

Tipologia e sociolinguistica: verso un approccio integrato allo studio della variazione

Atti del workshop della Società di Linguistica Italiana,
10 settembre 2020

a cura di

SILVIA BALLARÈ - GUGLIELMO INGLESE



S L I | Società di Linguistica Italiana

Tipologia e sociolinguistica: verso un approccio integrato allo studio della variazione

Atti del workshop della Società di Linguistica Italiana,
10 settembre 2020

a cura di

SILVIA BALLARÈ - GUGLIELMO INGLESE

Milano 2021

La Società di Linguistica Italiana (SLI), costituitasi a Roma nel 1967, ha lo scopo di promuovere studi e ricerche nel campo della linguistica, attraverso la creazione di una comunità di studiosi nel cui ambito trovi pieno riconoscimento e appoggio ogni prospettiva di ricerca linguistica teorica e applicata. La SLI tiene ogni anno un congresso internazionale di studi, e pubblica in volume alcuni dei contributi presentati al congresso. I manoscritti vengono valutati tramite un processo di revisione tra pari. Dal 2018 i volumi sono pubblicati con accesso libero a tutti gli interessati.

Iniziativa Dipartimenti di Eccellenza MIUR (L. 232 del 01/12/2016)



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
DIPARTIMENTO DI FILOLOGIA CLASSICA
E ITALIANISTICA

© 2021 SLI | Società di Linguistica Italiana
Roma
sito: www.societadilinguisticaitaliana.net



Edizione realizzata da
Officinaventuno
Via Flli Bazzaro, 18
20128 Milano - Italy
email: info@officinaventuno.com
sito: www.officinaventuno.com

ISBN edizione cartacea: 978-88-97657-47-7
ISBN edizione digitale: 978-88-97657-48-4

Indice

Presentazione	5
---------------	---

PARTE PRIMA

Introduzione

SILVIA BALLARÈ, GUGLIELMO INGLESE Sociolinguistica e tipologia: modelli e strumenti a confronto	11
--	----

PARTE SECONDA

Relazioni su invito

MASSIMO CERRUTI Variazione sociolinguistica e processi di grammaticalizzazione	53
NICOLA GRANDI Fattori sociolinguistici e costruzione del campione tipologico. Su alcune possibili interazioni tra tipologia e sociolinguistica	81

PARTE TERZA

Comunicazioni

MARCO ANGSTER Isolamento, cambio linguistico, integrità del sistema nel caso di una parlata in decadenza	103
FRANCESCA DI GARBO, ERI KASHIMA, RICARDO NAPOLEÃO DE SOUZA, KAIUS SINNEMÄKI Concepts and methods for integrating language typology and sociolinguistics	143
FABIO GASPARINI Nominalizzazione indipendente in sudarabico moderno	177
VÍCTOR LARA BERMEJO Evidential futures in Ibero-Romance	207

ANTONIETTA MARRA	
Forme di futuro nello Slavo del Molise tra sociolinguistica, tipologia e linguistica del contatto	225
EMANUELE MIOLA	
Varietà marginali di varietà substandard: appunti tipologici sui gerghi gallo-italici	249
ADRIANO MURELLI	
Strategie di relativizzazione nelle lingue europee: tra tipologia e sociolinguistica	269
Autrici e autori	301

Presentazione

Questo libro è una piccola anomalia rispetto agli altri volumi della collana in cui appare. Esso, infatti, non raccoglie gli atti del congresso della Società di Linguistica Italiana che, a causa dell'emergenza COVID-19, non ha potuto tenersi come da programma nel 2020. Questo volume contiene i contributi del workshop *Tipologia e sociolinguistica: verso un approccio integrato allo studio della variazione* che ha avuto luogo in modalità telematica il 10 settembre 2020.

Il tema scelto per il workshop mirava a portare a una riflessione sui rapporti tra fenomeni di variazione intralinguistica e interlinguistica. Le due prospettive, ovvero tipicamente quella della sociolinguistica variazionista e quella della linguistica tipologica, confluiscono di fatto solo di rado nella descrizione e nell'interpretazione di fenomeni di variazione, sebbene condividano un ampio terreno comune, avendo entrambe l'obiettivo di descrivere e analizzare la variazione linguistica.

Il volume è organizzato come segue. La prima parte, di cui fa parte anche questa breve presentazione, comprende un contributo dei due curatori, in cui si discutono in maniera sistematica affinità tra teorie e metodi della tipologia e della sociolinguistica. La seconda parte è dedicata ai due relatori invitati al workshop, Massimo Cerruti e Nicola Grandi. Il primo discute del rapporto tra sociolinguistica e studi di grammaticalizzazione, prendendo in esame alcuni processi in italiano e dialetti italo-romanzi e possibili casi di interferenza tra i due; il secondo invece offre una panoramica dei problemi che si pongono per il campionamento tipologico qualora si tenga conto anche di fattori sociolinguistici come il numero di parlanti per lingua.

La terza e ultima parte del volume è dedicata ai contributi degli altri partecipanti del workshop. Quello di Marco Angster riprende i temi dei due contributi dei relatori invitati e affronta il caso del *tisch* di Gressoney, una varietà alemanna parlata nell'Italia

nord-occidentale, mettendo in luce il ruolo del contatto per i fenomeni di mutamento linguistico osservabili in questa varietà e le conseguenze di questa particolare situazione per il campionamento tipologico. La tematica del contatto e del mutamento caratterizza anche il contributo di Francesca Di Garbo et al., in cui si presentano l'inquadramento teorico di un progetto volto a investigare sistematicamente il rapporto tra contesti sociolinguistici e diversità linguistica in comunità in contatto in una prospettiva tipologica. All'applicazione di modelli tipologici per lo studio di varietà linguistiche è dedicato il lavoro di Fabio Gasparini, che discute come la nozione tipologica di nominalizzazione indipendente sia utile per meglio comprendere i fenomeni di relativizzazione in due varietà del sudarabico moderno. In modo analogo, Víctor Lara Bermejo si serve della nozione di evidenzialità per investigare, anche grazie al ricorso a dati quantitativi, la funzione del futuro morfologico in diverse varietà non colte e dialettali iberoromanze. Costruzioni di futuro sono anche l'oggetto dell'articolo di Antonietta Marra, che passando in rassegna diverse forme di futuro nello Slavo del Molise, ne discute il rapporto con forme sia dell'originaria area balcanica sia delle varietà italo-romanze circostanti. Emanuele Miola affronta invece un altro tema proprio dell'approccio della tipologia sociolinguistica e investiga i processi di complessificazione e semplificazione in alcuni gerghi gallo-italici. Infine, il lavoro di Adriano Murelli offre una nuova prospettiva sulle strategie di relativizzazione in Europa, mostrando come l'inclusione di varietà standard e non-standard nello stesso campione contribuisca a raffinare ulteriormente la tipologia delle frasi relative nelle lingue d'Europa.

Prima di congedarci, desideriamo ringraziare Giuliano Bernini e Nicola Grandi per il sostegno datoci durante i lavori del workshop e per aver proposto di trasformare i contributi presentati in un volume pubblicato in questa collana. Cogliamo l'occasione per ringraziare anche i membri del comitato scientifico del workshop, gli autori, che hanno dato disponibilità a partecipare alla giornata di studi a distanza e poi a contribuire al volume, e infine tutti i revisori che hanno consentito di sottoporre tutti i capitoli qui presentati a un processo di *double blind peer review*.

Speriamo che i contributi raccolti nel presente volume possano costituire una solida base di partenza per future riflessioni circa il rapporto tra i vari rami della linguistica che si occupano di variazione.

Milano, giugno 2021

Silvia Ballarè e Guglielmo Inglese

PARTE PRIMA

INTRODUZIONE

SILVIA BALLARÈ, GUGLIELMO INGLESE

Sociolinguistica e tipologia: modelli e strumenti a confronto¹

La variazione linguistica è oggetto d'indagine privilegiato di diverse branche della linguistica, tra cui la sociolinguistica, che si occupa principalmente di variazione intralinguistica, e la tipologia, che studia invece la variazione interlinguistica. Nonostante le due discipline raramente collaborino nello studio dei fatti di variazione, esse hanno tuttavia sviluppato nel tempo una serie di modelli teorici e strumenti metodologici affini per l'analisi della variazione. In questo lavoro, ci poniamo come obiettivo di discutere alcuni di questi punti di contatto tra le due discipline, con lo scopo di gettare le basi per lo sviluppo di un approccio integrato allo studio della variazione.

Parole chiave: sociolinguistica, tipologia, variazione, mutamento.

1. Introduzione

La diversità e la variazione sono lo stato 'naturale' delle lingue (Grandi 2020). Almeno preliminarmente, possiamo intendere la variazione come l'esistenza di strategie formalmente diverse per esprimere un medesimo contenuto. Essa si presenta a diversi livelli: nell'idioletto di singoli parlanti, all'interno di comunità linguistiche, tra lingue distanti nello spazio e nel tempo. Tuttavia, la *natura* della variazione è la medesima: "nelle diverse lingue e nelle varietà di una stessa lingua sono in opera fondamentalmente le stesse tecniche di variazione, la

¹ La concezione e lo sviluppo del lavoro sono opera comune dei due autori. Per finalità accademiche, specifichiamo che Silvia Ballarè ha scritto i §§ 2.1, 2.2, 3.1, 4.2, 5.1 e Guglielmo Inglese i §§ 2.3, 3.2, 3.3, 4.1, 5.2. I §§ 1, 2, 3, 4, 5 e 6 sono stati scritti da entrambi gli autori.

Per quanto riguarda Silvia Ballarè, il contributo è frutto della ricerca svolta nell'ambito del progetto PRIN 2017 UniverS-Ita *L'italiano scritto degli studenti universitari: quadro sociolinguistico, tendenze tipologiche, implicazioni didattiche* (Settore ERC SH4, Prot. 2017 LAP429). Guglielmo Inglese ringrazia FWO – Research Foundation Flanders (grant. n. 12T5320N).

gamma di variazione che troviamo all'interno delle lingue costituisce spesso una scelta tra le tante possibilità che troviamo generalmente manifestate nei sistemi linguistici" (Berruto 2009: 23).

Considerata la sua pervasività come fatto linguistico, non stupisce perciò che la variazione sia tradizionalmente oggetto privilegiato di indagine di diverse branche della linguistica. In questo contributo, ci concentriamo su due discipline in particolare, la sociolinguistica e la linguistica tipologica o tipologia (v. già Kortmann 2004a).

Le due discipline si occupano apparentemente di ambiti distinti. La tipologia "concerns itself with the study of structural differences and similarities between languages" (Velupillai 2012: 15), mentre la sociolinguistica tipicamente studia "the correlation of dependent linguistic variables with independent social variables (Chambers 2003: ix)². La prima si basa sullo studio di campioni rappresentativi della varietà genealogica e geografica delle lingue del mondo (v. Grandi questo volume) mentre la seconda si occupa della variazione all'interno di una medesima lingua. In altre parole, la tipologia studia la variazione interlinguistica, mentre la sociolinguistica quella intralinguistica (o, nei termini di Grandi 2020, la variazione orizzontale e verticale)³.

In effetti, l'interesse per la variazione delle due discipline muove da premesse e scopi differenti. Semplificando, la tipologia si pone come obiettivo lo studio della diversità interlinguistica, con lo scopo di identificare i limiti della variazione⁴. Tramite uno studio empirico e su larga scala di cosa è e cosa non è possibile nelle lingue del mondo, la tipologia aspira a gettare luce sulla natura stessa del linguaggio umano, da una prospettiva che si pone in alternativa all'approccio tradizionalmente portato avanti dalla grammatica generativa (Daniel 2010). La sociolinguistica, invece, si occupa di studiare fatti linguistici che hanno un significato sociale al fine di capire "the mechanisms which link extra linguistic phenomena (the social and cultural) with

² Qui e nel seguito del lavoro, parlando di *sociolinguistica*, ci riferiamo principalmente al ramo della disciplina che si occupa dello studio della variazione linguistica.

³ Prescindiamo in questa sede dalla *vexata quaestio* relativa a lingua/dialetto e quindi a come distinguere la variazione intralinguistica (varietà diatopiche diverse) da quella interlinguistica (lingue diverse) (per una discussione recente rimandiamo a Gil 2016).

⁴ Interessante è notare come in tempi recenti la tipologia linguistica sia stata inclusa, insieme alla linguistica descrittiva, all'interno di una più generale *diversity linguistics* (es. Haspelmath 2014-).

patterned linguistic heterogeneity (the internal, variable, system of language)” (Tagliamonte 2012: XIV).

Un importante punto in comune tra la sociolinguistica e la tipologia è che entrambe le discipline si occupano di dimostrare come la variazione linguistica, a prescindere da dove la si osservi, non sia casuale, ma avvenga all’interno di limiti ben precisi e segua una particolare sistematicità. Questo interesse comune fa sì che, sebbene di rado le due discipline convergano nella descrizione e nell’analisi dei fenomeni di variazione, esse abbiano sviluppato nel tempo modelli teorici e strumenti per lo studio della variazione che mostrano notevoli affinità.

In questo contributo, ci poniamo l’obiettivo di illustrare alcuni dei principali punti di contatto che esistono tra la sociolinguistica e la tipologia, soprattutto in riferimento all’apparato teorico e metodologico, oltre che agli strumenti di ricerca. Questa breve rassegna ragionata, che non ha alcuna pretesa di esaustività, mira a favorire un maggiore dialogo tra le due discipline, auspicando di contribuire a gettare le basi per un approccio integrato allo studio della variazione linguistica.

Il contributo è strutturato come segue. In primo luogo, discutiamo di come la sociolinguistica e la tipologia abbiano sviluppato modelli teorici simili (§ 2) oltre che strumenti e metodi affini (§ 3) per affrontare problemi inerenti allo studio della variazione linguistica. La sezione 4 è dedicata al ruolo della diacronia nello studio della variazione inter- e intralinguistica. Nella sezione 5 offriamo una breve panoramica degli studi più recenti che mostrano come le due discipline abbiano fortemente beneficiato dell’influsso reciproco. Infine, la sezione 6 contiene alcune note conclusive.

2. *Modelli per lo studio della variazione*

La tipologia e la sociolinguistica hanno storicamente affrontato indipendentemente l’una dall’altra una serie di questioni metodologiche che mostrano in realtà numerosi punti di contatto. In particolare, entrambe le discipline hanno sviluppato dei modelli teorici per lo studio e l’analisi della variazione articolati in una serie di componenti che manifestano significative affinità. In questa sezione, ne discutiamo brevemente tre. Il primo ha a che fare con l’individuazione dei tratti linguistici da prendere in analisi e coinvolge dunque la discussione da un lato della nozione di *variabile sociolinguistica* e dall’altro di

categorie descrittive e concetti comparativi. Il secondo, invece, riguarda oggetti più complessi poiché costituiti dall'aggregazione di tratti linguistici, ovvero *varietà sociolinguistiche* e di *tipi linguistici*. Il terzo, infine, riguarda le relazioni che possono instaurarsi tra tratti linguistici e rimanda a *universali, gerarchie e scale implicazionali*.

2.1 Variabili sociolinguistiche e concetti comparativi

Una prima questione che ci si trova ad affrontare nel tentativo di descrivere la variazione – e dunque legata al concetto di variazione stessa – riguarda la necessità di delimitare il dominio dell'indagine. In altre parole, semplificando un poco, se l'obiettivo della ricerca è studiare come un certo fenomeno linguistico si manifesti in diverse (varietà di) lingue, sarà necessario elaborare un modello che permetta di individuare i confini del fenomeno stesso.

In ambito sociolinguistico, è cruciale la nozione di *variabile sociolinguistica* la cui definizione “is the first and also the last step in the analysis of variation” (Labov 2004: 7). Essa può essere intesa come un insieme di “alternative ways of saying “the same” thing” (Labov 1972: 188). Le realizzazioni concrete di questi “modi alternativi” sono dette *varianti* e tra esse deve esserci una relazione di “equivalenza semantica”, ovvero devono essere portatrici dello stesso significato. Perché la variabile assuma rilevanza sociolinguistica, le varianti devono correlare con fatti extralinguistici (come, ad esempio, la caratterizzazione sociale del parlante, la sua provenienza geografica o il grado di formalità del contesto, etc.). Un celebre esempio di variabile sociolinguistica riguarda la stratificazione sociale di (r) nel parlato di New York studiata da Labov (1966). Per studiare la variabile, vennero raccolte e analizzate diverse produzioni della stringa “fourth floor” in grandi magazzini destinati a clientela appartenente a diverse classi sociali. Considerando diversi parametri (tra cui, ad es., il costo della merce venduta), i tre grandi magazzini vennero ordinati dal più caro al più economico come segue: Saks, Macy's e S. Klein. Le realizzazioni concrete vennero poi ricondotte a due casi, ovvero r1 per “plainly constricted value of the variable” (Labov 1966: 173) e r0 per “unconstricted schwa, lengthened vowel, or no representation” (Labov 1966: 173). I risultati misero in luce il valore sociolinguistico della variabile: la variante r1 in modalità esclusiva si trova realizzata nei 3 grandi magazzini con i seguenti rapporti di frequenza: Saks

30/68 (0,44), Macy's 20/125 (0,16) e S. Klein 4/71 (0,06). Il risultato mise in luce dunque la caratterizzazione sociale della variabile: r1, infatti, si presenta molto più frequentemente nelle produzioni dei parlanti di classi sociali basse rispetto a quelle delle classi sociali più alte (0,44 di contro a 0,06).

Già a partire dagli anni '70 si aprì il dibattito circa i limiti dell'applicabilità del modello⁵. Esso, infatti, ben si presta a dare conto di fatti di variazione fonetico-fonologica ma le cose sono destinate a complicarsi non appena si coinvolgono altri piani di strutturazione linguistica. Salendo di livello, infatti, le unità di analisi sono *per definizione* portatrici di un significato e dunque implicano riflessioni ulteriori per quanto riguarda il principio di "equivalenza semantica". Sankoff (1972: 58) allarga le maglie della definizione di variabile sociolinguistica, scrivendo che "whenever there are options open to a speaker, we can infer from his or her behaviour an underlying set of probabilities". Lavandera, per includere tratti linguistici di altri livelli di analisi, parla di "functional comparability" (1978: 181) di forme in determinati contesti. Labov, prendendo in considerazione variabili morfosintattiche, limita l'equivalenza semantica al piano referenziale (Labov 1972: 271; 1978; Weiner & Labov 1983).

Nel corso degli anni, il raggio di azione di questo strumento è stato ampliato ed esso è stato impiegato anche per dare conto di fatti di variazione sociolinguistica relativi ad altri livelli di analisi, quali, ad esempio, morfologia, sintassi e lessico. Nel discutere questo ampliamento, Wolfram (1993: 195-196), facendo riferimento a caratteristiche strutturali⁶ delle varianti, scrive che si possono avere varianti riconducibili a una categoria strutturale, a fonemi, a elementi che occorrono in un certo contesto linguistico, a relazioni sintattiche, al posizionamento di un certo elemento rispetto ad un altro o a elementi lessicali. Conclude dunque che la variabile sociolinguistica "is founded in a linguistically-defined unit of some type, although this delineation is fairly broad, ranging from syntactic relationships to particular lexical items" (Wolfram 1993: 196).

⁵ Per una breve rassegna critica, v. Berruto (2007 [1995]: 139-145) e Tagliamonte (2006: 70-76).

⁶ Per una discussione attorno alla nozione di *identità di struttura* si rimanda a Cerruti (2011).

Il piano più problematico rimane certamente quello pragmatico⁷; le ragioni sono piuttosto evidenti se si considera l'astrattezza intrinseca delle categorie adottate per dare conto dei valori pragmatici (per una discussione su questo aspetto v. Pichler 2013: 6-9).

A questo punto, possiamo citare un aspetto fondamentale intrinsecamente legato alle variabili sociolinguistiche ovvero il "principle of accountability" (Labov 1982: 30), secondo cui per studiare il comportamento di una variabile è necessario dare conto delle occorrenze di *tutte* le varianti, così come delle non-occorrenze; per questo motivo, è necessario individuare i contesti linguistici in cui la variabile studiata potrebbe potenzialmente occorrere. Se questo obiettivo non risulta particolarmente problematico per fatti di fonetica, morfologia o sintassi, esso presenta certamente difficoltà per marcatori pragmatico-discorsivi.

Recentemente, almeno in relazione ad alcuni studi di caso, è stato proposto un rovesciamento del modello tradizionale della variabile sociolinguistica in modo tale da considerare un tipo di variazione associato ad un'unica realizzazione formale (per cui v. § 4.2).

La linguistica tipologica è essenzialmente una disciplina comparativa, che si occupa della categorizzazione di fatti (inter-)linguistici (Moravcsik 2016). La necessità di delimitare il campo di indagine della variazione, cioè garantire la comparabilità e la corretta categorizzazione, è perciò il prerequisito fondamentale dello studio tipologico⁸. In altre parole, "cross-linguistic identification of studied phenomena presupposes a procedure which ensures that we compare like with like" (Koptjevskaja-Tamm *et al.* 2015: 436). Il problema è relativamente di facile soluzione nel caso di alcuni domini più basilari della tipologia lessicale (Koptjevskaja-Tamm 2012): è il caso del lessico del colore, che può essere investigato somministrando a parlanti di lingue diverse uno stesso stimolo visivo con diversi colori a cui attribuire il nome corrispondente, come nel classico studio di Berlin &

⁷ Tra i diversi studi possiamo citare almeno Pichler (2013), Tagliamonte (2005), Schneider & Barron (2008) e Cameron & Schwenker (2013). Per una recente rassegna si rimanda a Sansò (2020: 81-104).

⁸ Per una panoramica di veda Evans (2020). Due numeri monografici di *Linguistic Typology*, 20(2) (2016) e 24(3) (2020), e un volume edito da Alfieri *et al.* (2021) sono stati recentemente dedicati interamente al problema della comparabilità in tipologia.

Kay (1969)⁹. È invece meno immediato garantire la comparabilità tra costruzioni a livello morfosintattico. Volendo ad esempio investigare il comportamento degli aggettivi nelle lingue del mondo, è necessario stabilire una definizione della nozione di aggettivo che sia ugualmente applicabile a lingue strutturalmente anche molto distanti. Il problema è perciò quali criteri vadano inclusi in una definizione interlinguistica, soprattutto tenendo in considerazione che l'applicazione dei criteri utilizzati per individuare la classe degli aggettivi in lingue come l'italiano non consente di individuare una classe equivalente di aggettivi in lingue come lo yurok, dove concetti tipicamente espressi da aggettivi in italiano come 'veloce' sono espressi da costruzioni che mostrano le stesse caratteristiche delle costruzioni identificabili come verbi (Dixon 1982). Si tratta di un problema che coinvolge praticamente ogni fenomeno morfosintattico, dal momento che molte delle nozioni della tradizione grammaticale occidentale non sono immediatamente applicabili al livello interlinguistico. Si pensi alle nozioni di soggetto e oggetto (cf. Bickel 2010; Witzlack-Makarevich 2019) e frase subordinata (Cristofaro 2003), fino alla nozione stessa di parola (Haspelmath 2011).

Una possibile soluzione, che consente di prescindere da criteri formali e strutturali specifici di singole lingue, è quella di fondare la comparazione interlinguistica esclusivamente su concetti semantico-funzionali universali (cfr. Stassen 1985: 14; Croft 2003: 13-19). Queste definizioni universali e astratte sono state definite da Haspelmath (2010, 2018) *concetti comparativi* (da contrapporsi alle *categorie descrittive* proprie di singole lingue). I concetti comparativi sono lo strumento essenziale che consente alla tipologia di operare comparazioni e generalizzazioni tra fenomeni simili in lingue diverse, dal livello della fonologia a quello della semantica. Per tornare all'esempio degli aggettivi, la comparabilità tra costruzioni diverse si può basare sul concetto comparativo di aggettivo come "a lexeme that denotes a descriptive property and that can be used to narrow the reference of a noun" (Haspelmath 2010: 670)¹⁰.

⁹ In realtà anche questo approccio non è privo di problemi, specialmente se applicato a domini del lessico più astratti, come discusso nel dettaglio da Koptjevskaja-Tamm *et al.* (2015). Sull'utilizzo di stimoli non linguistici v. anche § 3.2.

¹⁰ Sulla tipologia degli aggettivi in particolare v. Alfieri (2014) e Rießler (2016: 6-7).

In questo senso, la tipologia fa eco alla nozione di “equivalenza semantica” discusso sopra per la sociolinguistica, nel senso che nel caso dei concetti comparativi al pari di quello delle variabili sociolinguistiche, il problema della comparabilità è stato affrontato analogamente dando sistematica priorità a parametri di tipo semantico-funzionale nella delimitazione del campo di indagine. Tuttavia, pur senza mettere in discussione la priorità del criterio funzionale, studi più recenti hanno messo in luce come anche proprietà formali/strutturali possano essere tenute in considerazione nella formulazione di concetti comparativi (es. Miestamo 2005: 39-45; Stassen 2010; Croft 2016), purché anch’esse prescindano da categorie descrittive di lingue particolari e siano basate su altri concetti comparativi.

Per loro natura, i concetti comparativi rappresentano essenzialmente un approccio di tipo onomasiologico alla variazione. Ciò non vuol dire che la tipologia sia del tutto estranea a una prospettiva semasiologica. Studi di tipo semasiologico sono stati svolti in particolare nell’ambito della tipologia lessicale, in cui una particolare attenzione è stata rivolta ai diversi significati che possono essere associati a singole parole, cioè ai pattern di polisemia e colessificazione in una prospettiva interlinguistica (Koptjevskaja-Tamm 2012: 347; v. anche François 2008; List *et al.* 2018). Oltre che al livello puramente lessicale, l’approccio semasiologico è anche alla base del modello delle mappe semantiche. Esse sono state introdotte nella linguistica tipologica con lo scopo di rappresentare visivamente le connessioni che esistono tra significati espressi da costruzioni linguistiche, inclusi anche significati più astratti di tipo grammaticale (rimandiamo a Croft 2003: 133-142; Haspelmath 2003; Cristofaro 2010; Cysouw 2007, 2010; van der Auwera 2013; Georgakopoulos & Polis 2018 per una discussione dettagliata). Quello che ci preme sottolineare qui è come i vari significati rappresentati su una mappa semantica non siano stabiliti aprioristicamente, ma emergano direttamente dal confronto dei pattern di polifunzionalità di specifiche costruzioni in lingue diverse.

2.2 Varietà sociolinguistica e tipo linguistico

Come messo in luce da Grandi in una serie di pubblicazioni (2018, 2020), una caratteristica comune del metodo di analisi linguistica della sociolinguistica e della tipologia è che, una volta individuato il dominio della variazione, come discusso in 2.1, i singoli tratti presi in

esame possono essere organizzati in aggregati che rendono conto della sistematicità della variazione. Le due nozioni in questione sono quella di varietà sociolinguistica e di tipo linguistico.

La nozione di *varietà di lingua* venne introdotta per discutere il fatto che la variazione osservabile nelle lingue non è casuale; le lingue, anzi, andrebbero viste “as an object possessing orderly heterogeneity” (Weinreich *et al.* 1968: 100). Possiamo intendere una varietà di lingua come “la realizzazione del sistema linguistico in, o meglio presso, classi di utenti e di usi” (Berruto 2007 [1995]: 63; sull’argomento v. anche Hudson 1998 [1980]: 27-76 *inter al.*). Una varietà di lingua, dunque, è costituita da un insieme di tratti linguistici – ai vari livelli di strutturazione – che tendono a co-occorrere presso certi parlanti, in certe situazioni d’uso. È importante sottolineare che “i tratti linguistici tipici di una varietà devono inoltre essere congruenti, vale a dire dotati di un certo grado di omogeneità strutturale, che fa sì che obbediscano a specifiche regole di co-occorrenza” (Berruto 1995 [2007]: 64); vale a dire che perché un certo enunciato sia *coerente* dal punto di vista sociolinguistico esso deve essere formato da item riconducibili alla stessa varietà.

I tratti linguistici possono occorrere con frequenza diversa in varietà linguistiche diverse oppure possono essere caratteristici solo di una certa varietà; a mo’ di esempio, in quest’ottica, può essere utile considerare la discussione attorno alla relazione tra italiano popolare e italiano colloquiale condotta da Berruto (2012: 139-143) in termini di frequenza di alcuni tratti attestati in entrambe le varietà. Osservando questo aspetto da un’altra prospettiva, possiamo aggiungere che le diverse varietà linguistiche non sono da concepirsi come oggetti discreti. Esse, infatti, si possono disporre su un continuum, creando un “ordered set of elements arranged in such a way that between two adjacent entities of the set (in this case, language varieties) there are no sharp boundaries, but rather a gradual, fuzzy differentiation, each variety sharing some sociolinguistically marked features with adjacent varieties” (Berruto 2010: 235).

Questa visione della lingua, di chiara matrice strutturalista, come elemento “organizzato” in varietà è stata oggetto di numerose critiche negli ultimi anni mirate a mettere in discussione l’individuabilità delle varietà linguistiche e/o la loro stessa esistenza (ma cfr. Hinskens & Guy 2016; per una discussione, si rimanda a Ghyselen & De Vogelaer

2018). Tuttavia, studi recenti, resi possibili anche dalla presenza di corpora corredati da metadati e dall'applicazione di metodi quantitativi all'analisi linguistica, hanno permesso di dimostrare la presenza concreta delle varietà linguistiche. Prendendo il caso dell'italiano, possiamo citare lo studio condotto da Cerruti & Vietti (in stampa) in cui, considerando una parte del corpus KIParla (Mauri *et al.* 2019) si valuta la distribuzione di tratti del "vecchio" e del "nuovo" standard in produzioni con diversa caratterizzazione diafasica. L'analisi dei componenti principali fa emergere che i due set di tratti si comportano in maniera coerente: i tratti del vecchio standard occorrono in maniera compatta in produzioni altamente formali e, specularmente, quelli del "nuovo" standard occorrono coerentemente in produzioni più informali.

La nozione di tipo linguistico risale ai primordi della tipologia linguistica, e si basa sull'assunto che la diversità delle lingue del mondo può essere ricondotta a una serie ridotta di tipi generali diversi (cfr. Ramat 2010; sulla nozione di tipo in tipologia e in altre scienze si veda in particolare Round & Corbett 2020). In effetti, la tipologia stessa può essere intesa come "the study and interpretation of linguistic or language types" (Velupillai 2012: 15). Esempi di tipi morfologici tradizionali sono le lingue isolanti, agglutinanti, incorporanti e fusive. Al livello morfosintattico una tipologia per tipi prevede la distinzione tra lingue con diversi tipi di allineamento, come quello nominativo-accusativo o quello ergativo-assolutivo. Infine, al livello dell'ordine dei costituenti due tipi principali sono postulati, cioè a testa iniziale e a testa finale. Una importante caratteristica dei tipi linguistici, come anche sottolineato da Grandi (2014: 11-15), è che questi non costituiscono semplicemente aggregati casuali di caratteristiche, ma piuttosto hanno un "carattere prettamente strutturale" (Grandi 2014: 14). In particolare, il tipo linguistico è significativo perché spesso consente di fare predizioni sulla (co-)occorrenza di particolari caratteristiche (v. anche § 2.3). Per esempio, come discusso da Greenberg (1963), sapere che una lingua ha un ordine basico O(ggetto) V(erbo) consente anche di prevedere che molto probabilmente la stessa lingua avrà postposizioni e costruzioni in cui il genitivo precede il nome che modifica (v. Song 2010).

È importante, tuttavia, ribadire come l'idea che i tipi linguistici offrano una categorizzazione olistica e discreta delle lingue, nella misura in cui una lingua aderisce pienamente a un certo tipo e i tipi siano nettamente distinguibili l'uno dall'altro, sia in realtà stata in buona parte abbandonata dalla ricerca tipologica più recente (Croft 2003: 42-45). Il tipo linguistico va infatti considerato come un costrutto ideale, che non trova la sua piena realizzazione in nessuna lingua naturale (Coseriu 1973: 253). Per citare un esempio, già a partire da Sapir (1921) l'utilità della classificazione delle lingue in isolanti *vs.* flessive è stata messa in dubbio ed è stata sostituita da una misurazione più accurata dell'indice di sintesi/analisi delle singole lingue, che risulta, tra le altre cose, dalla somma del numero di costruzioni analitiche e sintetiche di ciascuna lingua (v. Greenberg 1960; Siegel *et al.* 2014). Anche nel caso dell'allineamento è stato da tempo dimostrato come questa non sia una proprietà olistica, ma che invece l'individuazione delle relazioni grammaticali, e il loro raggruppamento, vada operata al livello delle singole costruzioni: non è perciò la lingua ad essere, ad esempio, nominativo-accusativa, ma piuttosto una particolare costruzione all'interno della lingua (v. Bickel 2010; Witzlack-Makarevich 2019 con riferimenti).

Da queste considerazioni, emerge chiaramente come la variazione interlinguistica in molti casi non sia organizzabile in un inventario di pochi tipi ben definiti in modo discreto, ma sia piuttosto da intendere come uno spazio multidimensionale, formato da diversi parametri, in cui le lingue si possono disporre in modo diverso (v. Arkadiev & Klamer 2018: 444 con particolare riferimento alla tipologia morfologica). Un'importante spinta in questa direzione è stata data dall'apertura della tipologia a studi di tipo quantitativo basati su grandi quantità di dati naturalistici (v. § 3.2.3). Questi studi hanno prodotto risultati particolarmente significativi nello studio dell'ordine dei costituenti, consentendo di verificare empiricamente l'idea che, ad esempio, le lingue non seguono necessariamente un singolo ordine di soggetto, oggetto e verbo, ma che diversi ordini possono coesistere in diversa misura nelle singole lingue (Levshina 2019; Gerdes *et al.* 2021). In questo senso, i tipi linguistici sono particolarmente affini alle varietà linguistiche, dal momento che anche questi sono organizzati in *continua*.

2.3 Relazioni implicazionali: universali, gerarchie e scale

Uno degli aspetti comuni forse più interessanti della sociolinguistica e della tipologia è lo sviluppo di potenti strumenti teorici che consentono di catturare l'esistenza di rapporti sistematici tra tratti, attestati in lingue o varietà di lingua diverse.

Uno dei fautori dell'introduzione delle "relazioni implicazionali" nella linguistica moderna fu Greenberg (1963). Esse possono essere espresse in termini di universali e gerarchie implicazionali.

La formulazione di universali linguistici è uno dei risultati maggiori dello studio della variazione interlinguistica: è tramite gli universali che si possono identificare quali sono i limiti della diversità delle lingue del mondo e predire quali tipi di lingue sono possibili e quali no (Croft 2003: Cap. 3; Moravcsik 2010). Gli universali sono di due tipi: assoluti e implicazionali, come mostrato in (1a-b) (Croft 2003: 52-53).

- (1) a. *Tutte le lingue hanno vocali orali*
 b. *Se una lingua ha vocali nasali, avrà anche nasali orali ma non viceversa*

L'universale in (1a) è universale perché descrive una proprietà condivisa da tutte le lingue del mondo. L'universale in (1b) è invece implicazionale perché mette in relazione tra loro almeno due tratti, rispetto ai quali l'occorrenza del primo implica il secondo ma non viceversa. È evidente come gli universali implicazionali siano teoricamente più interessanti degli universali assoluti, dal momento che solo i primi consentono di delimitare la variazione linguistica in modo tale da escludere l'esistenza di alcuni tipi linguistici. Ad esempio, l'universale in (1b) esclude la possibilità dell'esistenza di lingue che hanno solo vocali nasali, ma non orali.

Una gerarchia implicazionale è una catena mono-dimensionale di universali implicazionali, di cui l'*implicans* dell'uno costituisce l'*implicatum* dell'altro (Croft 2003: 122). Le gerarchie implicazionali sono strumenti teorici particolarmente utili per catturare la sistematica relazione tra un numero cospicuo di tratti e sono state formulate in riferimento a diversi domini lessicali e morfosintattici (Croft 2003: Cap. 5; Corbett 2010). Importanti gerarchie sono ad esempio la gerarchia di accessibilità alla relativizzazione, formulata da Keenan & Comrie (1977, 1979), e la gerarchia di animatezza o referenzialità (es.

Corbett 2000; Cristofaro 2013). In ambito lessicale, possiamo menzionare la gerarchia dei verbi di percezione (Viberg 1984) nonché le scale di transitività (Tsunoda 1985; Malchukov 2005) e spontaneità (Haspelmath 1993, 2016) dei lessemi verbali. Infine, come osserva Croft (2003: 153), le mappe semantiche (v. § 2.1) possono essere considerate un caso particolare di gerarchia implicazionale complessa, cioè bi-dimensionale.

Nella tradizione dell'approccio tipologico-funzionale, i rapporti tra diverse proprietà linguistiche catturati dagli universali implicazionali e dalle gerarchie non sono considerati casuali, ma sono tipicamente spiegati come il risultato di una serie di principi funzionali-cognitivi soggiacenti, come ad esempio il principio di economia, l'iconicità, la marcatezza, la memoria, la capacità di processing etc. (Moravcsik 2010; Sansò 2018).

In realtà, la ricerca tipologica più recente ha messo, almeno in parte, in dubbio la validità di questo tipo di spiegazioni sincroniche funzionali. In primo luogo, è stato proposto che la frequenza, connessa alla prevedibilità e all'efficienza di codifica, giochi un ruolo fondamentale nel predire correttamente alcuni universali nella marcatura asimmetrica di particolari valori di certe categorie grammaticali (ad es., le marche di plurale universalmente tendono ad essere espresse con un numero uguale o superiore di morfemi rispetto alle marche di singolare, e mai viceversa; si rimanda a Haspelmath 2006, 2008, 2021). Un altro tipo di critica viene invece dagli studi di tipologia diacronica, su cui ritorniamo nella Sezione 4.1.

In ambito sociolinguistico, le *scale implicazionali* sono tradizionalmente considerate uno strumento alternativo rispetto alle *regole variabili*, intese come espressione del comportamento di una variabile (v. Rickford 2002 *inter al.*). In realtà, *scale implicazionali* e *regole variabili* possono essere (e sono state) usate in combinazione, anche perché rappresentano oggetti diversi. Le *scale*, infatti, danno conto di rapporti che possono istituirsi tra variabili (Berruto 2007 [1995]: 156). Esse costituiscono uno strumento per ordinare secondo gerarchie una serie di tratti linguistici, verificando se la presenza di un certo tratto implichi la presenza di altri tratti in determinate varietà sociolinguistiche – o produzioni di parlanti con caratteristiche sociali note (per una discussione v. Berruto & Cerruti 2015: 131-136). Concretamente, la

scala è costituita da una matrice a doppia entrata in cui si collocano da una parte i tratti linguistici e dall'altra i parlanti o varietà di lingua¹¹. Le celle possono essere riempite con valori binari (ad es. +/-) che danno conto della presenza/assenza di un certo tratto in una certa produzione, oppure si possono avere intervalli di frequenza relativi all'occorrenza di un certo tratto in un certo tipo di produzioni (per un esempio v. Cerruti 2009: 235-254).

3. *Strumenti*

Già si è detto di come la sociolinguistica e la tipologia affrontino il tema della diversità linguistica adottando apparati teorici che mostrano notevoli punti di contatto (v. § 2). Non sorprende quindi che le due discipline si siano dotate nel tempo anche di una serie di strumenti di ricerca simili. In questa sezione intendiamo mettere in luce le affinità tra questi strumenti, non tanto con lo scopo di offrirne una descrizione dettagliata, ma piuttosto di evidenziare come questi strumenti siano nati, sebbene indipendentemente, come risposta alla necessità di affrontare problemi teorici e metodologici di natura simile.

3.1 Atlanti linguistici

Gli atlanti linguistici costituiscono un importante strumento per osservare la variazione linguistica nello spazio geografico. Nella tradizione dialettologica, almeno in una prima fase, essi sono stati dedicati soprattutto a fatti di variazione fonetica e lessicale e hanno trovato terreno fertile in paesi che presentavano grande eterogeneità linguistica, come Germania, Svizzera, Italia e Russia (Lameli 2010: 569) – si pensi, in ambito italiano, all' AIS (Jaberg & Jud 1928-1940) e all' ALI (Bartoli 1995-). Gli atlanti linguistici moderni sono costituiti da singole carte dedicate a un certo fenomeno in cui si riporta come questo fenomeno si realizza nei punti d'inchiesta nell'area considerata¹².

Più recentemente, sono stati realizzati atlanti tipologici sia (più spiccatamente) sociolinguistici. In ambito sociolinguistico, recente-

¹¹ La scala implicazionale può essere utilizzata anche per altri fini come ad es. la rappresentazione dei rapporti interni alle varietà del repertorio (per una breve rassegna v. Cerruti 2010: 29-30).

¹² Sull'argomento v. Chambers & Trudgill (2004 [1998]) e Lameli *et al.* (2011).

mente sono stati sviluppati atlanti linguistici per dare conto di fatti di variazione intralinguistica. In questi casi la singola carta potrebbe essere considerata come una variabile sociolinguistica e le diverse realizzazioni come le diverse varianti geografiche. Citiamo, a titolo meramente esemplificativo, il SAND – *Syntactische Atlas van de Nederlandse Dialecten* (Barbiers *et al.* 2005) e l'ANAE – *Atlas of North American English* (Labov *et al.* 2006). Nell'ottica di discutere punti di contatto tra tipologia e sociolinguistica, qualche parola di più merita forse l'eWAVE – *electronic World Atlas of Varieties of English* (Kortmann & Lunkenheimer 2013, versione digitale Kortmann *et al.* 2020) che, caratterizzato da una struttura del tutto analoga a quella del WALS – *Word Atlas of Language Structure* (Haspelmath *et al.* 2005, versione digitale Dryer & Haspelmath 2013, per cui v. oltre), costituisce un database di strutture morfosintattiche che caratterizzano 77 varietà geografiche di inglese. In questo caso, tuttavia, si ha una differenza significativa. Mentre nel WALS la variazione intralinguistica non viene considerata e ad ogni lingua è associata una sola variante della variabile considerata¹³, in questo caso ogni carta è associata ad una singola struttura (ad es. carta 154 *Multiple negation / negative concord*) e i valori associati alle diverse varietà di inglese sono espressi in termini di frequenza (da *feature is pervasive or obligatory* a *attested absence of feature*).

L'attenzione per la dimensione geografica e areale costituisce in qualche misura un interesse più recente della tipologia. L'atlante tipologico principale è senz'altro il WALS. I 144 capitoli del WALS forniscono informazioni sulla distribuzione di tipi linguistici per svariati domini, dalla fonologia alla sintassi, e ogni capitolo è corredato da una mappa che consente di visualizzare in modo intuitivo la distribuzione geografica dei vari tipi. Dal punto di vista metodologico, è importante notare come il WALS sia essenzialmente stato costruito seguendo la procedura più tradizionale della tipologia *type-based* (cf. Levshina 2019), nella misura in cui i vari capitoli sono prevalentemente basati sullo spoglio delle grammatiche di riferimento delle lingue in questione e le singole lingue sono classificate come appartenenti a un tipo o

¹³ Ad esempio, nella carta 122A: *Relativization on Subject* (Comrie & Kuteva 2013) l'italiano presenta il solo valore *relative pronoun* che dà conto della sola strategia con *il quale* ma trascura quella con *che*.

a un altro sulla base della occorrenza (o meno) di un particolare tipo di costruzione.

Il WALS ha anche agito da modello per la creazione di nuove risorse dall'impostazione simile. Citiamo qui per esempio l'APiCS – *Atlas of Pidgin and Creole Language Structures* (Michaelis *et al.* 2013), che sul modello del WALS contiene dati sulla distribuzione geografica di 130 tratti linguistici dei pidgin e delle lingue creole, e il PHOIBLE, che raccoglie dati sugli inventari fonologici di più di 2100 lingue visualizzabili anche su base geografica (Moran & McCloy 2019). Esistono anche atlanti 'tipologici' ristretti a particolari aree geografiche, come ad esempio SAILS – *South American Indigenous Language Structures* (Muysken *et al.* 2016) un database di tratti grammaticali di lingue del Sud America. Una organizzazione areale è anche alla base del *Diachronic Atlas of Comparative Linguistics* (DiACL; Carling 2017), che raccoglie dati sulla distribuzione di tratti linguistici in lingue antiche e moderne di tre macro-aree: Eurasia, Pacifico e Amazzonia. Infine, un progetto simile in spirito al WALS è *Grambank*, sviluppato all'interno del progetto *Glottobank*¹⁴, il cui obiettivo è di sistematizzare in un database, contenente anche informazioni di tipo geografico, dati su 195 tratti linguistici nelle lingue del mondo attualmente documentate.

3.2 I questionari

Un secondo tipo importante di strumento, in parte legato agli atlanti, è dato dai questionari. L'uso dei questionari nasce dalla comune esigenza della tipologia e della sociolinguistica di affrontare il problema della comparabilità (§ 2.1). I questionari solitamente si compongono di due parti. Una parte prettamente linguistica riguarda il dato linguistico da elicitare: questo viene solitamente fatto somministrando agli informanti un particolare stimolo, solitamente una parola o frase da tradurre nella lingua oggetto d'indagine. La seconda parte raccoglie invece una serie di informazioni accessorie al dato.

In ambito sociolinguistico il questionario è uno strumento largamente utilizzato per la raccolta di diversi tipi di dati; ad esempio, giudizi di accettabilità, informazioni circa repertori, usi linguistici e attitudini (v. ad es. Meyerhoff *et al.* 2015: 71-73). Queste informazioni,

¹⁴ glottobank.org.

di carattere linguistico, tipicamente sono corredate da una sezione del questionario dedicato ai metadati dell'informante (ad es. età, regione di origine, titolo di studio, ...) in modo da poter permettere in fase di analisi di verificare la presenza di eventuali correlazioni.

Quando l'obiettivo dell'indagine è, ad esempio, verificare la distribuzione o la caratterizzazione sociolinguistica di un certo tratto linguistico, il questionario può rivelarsi molto utile, soprattutto nel caso in cui non si disponga di corpora relativi alle varietà che siamo interessati a studiare. Il questionario, infatti, può essere somministrato ad un campione di informanti creato *ad hoc* in base alle esigenze della nostra ricerca. Inoltre, esso permette l'elicitazione mirata di tratti linguistici che, nelle produzioni spontanee, potrebbero presentarsi poco frequentemente. Tuttavia, il questionario presenta anche diversi limiti. In prima istanza, è opportuno sottolineare che i dati raccolti tramite esso costituiscono *valutazioni* espresse dall'informante, che possono anche di molto discostarsi dal suo uso concreto della lingua (v. Iannàccaro 2000 e più in generale per una riflessione metodologica v. Sanga 1991).

In aggiunta alla metodologia basata sullo spoglio di grammatiche, la tipologia si serve anche di questionari ideati con lo scopo di raccogliere dati uniformi su particolari fenomeni tra le lingue del mondo. Forse il più noto questionario tipologico è quello sviluppato da Östen Dahl per il suo lavoro sulla tipologia delle marche tempo-modo-aspettuali (Dahl 1985), contenente un set di poco meno di 200 frasi stimolo¹⁵.

Il vantaggio metodologico dei questionari è che questi consentono di superare alcuni limiti inerenti al metodo *type-based* dello spoglio di grammatiche. In particolare, elicitando direttamente il dato tramite uno stimolo standardizzato, si minimizza il rischio di ignorare tratti presenti in una lingua ma che per qualche motivo non sono riconoscibili o documentati nella grammatica di riferimento. Va inoltre menzionato come, in tempi più recenti, oltre ai questionari contenenti stimoli linguistici sono anche stati sviluppati set di stimoli non linguistici di natura audio-visuale, per specifici domini sia morfosintattici

¹⁵ Per una lista di questionari tipologici si rinvia a <https://www.eva.mpg.de/lingua/tools-at-lingboard/questionnaires.php>.

(es. costruzioni reciproche, Evans et al. 2004) sia lessicali (es. colori, Berlin & Kay 1969, eventi di moto, Bohnemeyer & Caelen 1999)¹⁶.

3.3 I corpora

I corpora linguistici sono risorse tradizionalmente usate per ricerche che vogliono concentrarsi sullo studio della variazione (socio)linguistica. Essi, infatti, permettono l'accesso a una mole considerevole di testi, spesso accompagnati da metadati relativi al contesto in cui sono stati raccolti e/o al parlante/scrivente coinvolto. Queste risorse sono particolarmente preziose anche perché permettono analisi di tipo quantitativo, che sin dalle origini hanno caratterizzato buona parte degli studi sociolinguistici. Molti sono i sociolinguisti che, nel corso degli ultimi decenni, si sono dedicati alla creazione di corpora e alla discussione relativa alle metodologie per la raccolta dati (v. ad es. Tagliamonte 2006) e alla rilevanza dei corpora per l'analisi sociolinguistica (Poplack 2021). In ambito italiano, non mancano risorse che permettono lo studio della variazione (per una rassegna v. Gorla & Mauri 2018) anche se è piuttosto recente la creazione di corpora corredati da un numero consistente di metadati, per cui si pensi ad esempio al corpus KIParla (Mauri *et al.* 2019) e a quello, attualmente in fase di allestimento, del PRIN 2017 *UniverS-ITA*¹⁷.

I corpora costituiscono invece un'acquisizione relativamente recente per la tipologia. Un'eccellente panoramica sull'uso dei corpora nella ricerca tipologica si trova in Levshina (2020), a cui rimandiamo per una discussione dettagliata dei tipi di corpora, degli interrogativi della ricerca tipologica investigabili tramite corpora, e dei vantaggi e limiti di una tipologia *corpus-based*.

In primo luogo, bisogna osservare come l'uso dei corpora per la tipologia sia metodologicamente più complesso dell'uso dei corpora per la sociolinguistica. Per consentire studi tipologici basati su corpora, infatti, è necessario avere a disposizione una sufficiente quantità di corpora di lingue diverse che siano sufficientemente simili però da consentire una comparazione proficua. Due sono le soluzioni principali adottate in questo senso: l'uso di corpora paralleli, cioè di tradu-

¹⁶ Rimandiamo a Koptjevskaja-Tamm et al. (2015: 441-442) per una rassegna.

¹⁷ site.unibo.it/univers-ita.

zioni allineate di un medesimo testo in più lingue (Cysouw & Wälchi 2007), o, più recentemente, l'uso di corpora nativi che prevedono uno stile di annotazione simile, come nel caso del progetto di Universal Dependencies (Zeman *et al.* 2020). Un importante progetto in corso, che ha come scopo lo sviluppo di corpora di parlato di oltre 150 lingue (con particolare attenzione a lingue a rischio e/o scarsamente documentate) è il DoReCO project (Paschen *et al.* 2020).

La dimensione quantitativa non è nuova alla tipologia (v. ad esempio Greenberg 1960 sulla complessità morfologica), ma, come osservato da Levshina (2020), l'uso di corpora apre a una nuova era di una tipologia quantitativa basata su una ben più ricca base empirica di dati linguistici reali. È bene notare che l'uso di corpora non costituisce semplicemente di un'accortezza metodologica, ma ha anche importanti ripercussioni sul piano teorico. In particolare, i corpora consentono di superare la visione spesso categorica dei tratti linguistici che emerge dal metodo tradizionale dello spoglio grammaticale (per cui una particolare lingua è codificata sulla base del suo mostrare o meno un certo valore per un tratto linguistico) a favore di una visione che meglio rifletta la natura graduale dei fenomeni linguistici. Ciò significa, ad esempio, rivedere la nozione di tipo linguistico categorico in favore di una descrizione più empiricamente accurata dei fatti di ciascuna lingua (§. 2.2). Un campo in cui un approccio quantitativo ha dato particolari frutti è, ad esempio, lo studio interlinguistico dell'ordine delle parole. Studi recenti di Levshina (2019) e Gerdes *et al.* (2021), hanno confermato empiricamente l'idea che le lingue non seguono necessariamente un singolo ordine di soggetto, oggetto e verbo, cioè non corrispondano pienamente a un singolo tipo es. SVO, ma che diversi ordini possono coesistere in diversa misura nelle singole lingue (Gerdes *et al.* 2021). In questo contesto, ci preme notare che, come sottolineato da Levshina (2020: 8), nel momento in cui lo studio tipologico si interessa della natura graduale del linguaggio e dell'interazione tra la realizzazione di particolari tratti e il contesto in cui occorrono, la tipologia diventa a tutti gli effetti uno studio di tipo variazionista, che può quindi beneficiare dei metodi della tradizione sociolinguistica.

4. *Mutamento e variazione: il ruolo della diacronia*

Nella tradizione strutturalista, sincronia e diacronia sono due prospettive dell'analisi linguistica incompatibili tra loro (Saussure 2016 [1916]: 174). In questa prospettiva, la variazione (intra- e interlinguistica) è essenzialmente un fatto di sincronia, mentre il mutamento pertiene alla dimensione diacronica. In realtà, la dicotomia tra sincronia e diacronia è in parte fuorviante, e le due dimensioni vanno piuttosto pensate come “due facce della stessa medaglia” (Seiler 2018: 82; v. anche Giacalone Ramat *et al.* 2012). Il rapporto tra le due dimensioni è infatti duplice (v. es. Croft 2000; Luraghi 2010; Seiler 2018): da un lato, la variazione è il risultato del mutamento, nella misura in cui l'esistenza di diverse varianti presuppone l'introduzione, tramite innovazione interna o esterna, di una nuova variante in un particolare sistema; dall'altro, la variazione sincronica è la base del mutamento inteso come la selezione e convenzionalizzazione di una specifica strategia. Un grande contributo alla “dinamicizzazione” dei rapporti tra sincronia e diacronia si deve agli studi sui processi di grammaticalizzazione (v. Hopper & Traugott 2003 e Lehmann 2015 *inter al.*). I processi di grammaticalizzazione tipicamente coinvolgono elementi lessicali che, talvolta tramite una serie di fasi intermedie ben precise, assumono progressivamente un valore grammaticale perdendo allo stesso tempo autonomia morfosintattica. Questi processi diacronici si avviano nell'uso concreto, *on-line*, della lingua grazie alla reinterpretazione operata dai parlanti di costruzioni di partenza in costruzioni di arrivo (v. oltre).

In questa sezione, discutiamo brevemente come la tipologia e la sociolinguistica abbiano integrato la dimensione diacronica all'interno di studi di variazione inter- e intralinguistica.

4.1 Tipologia diacronica

Fin dalle origini della ricerca tipologica la dimensione sincronica e quella diacronica sono state strettamente intrecciate. Si prenda il caso della tipologia morfologica: già Humboldt, nel formulare la nota classificazione delle lingue in isolanti, agglutinanti, fusive e incorporanti, non solo propone l'esistenza di quattro tipi sincronicamente distinti, ma stabilisce anche una connessione diacronica tra questi, nel senso che i quattro tipi da considerarsi quattro stadi successivi in un ideale processo evolutivo.

Anche in tempi più recenti, la necessità di una tipologia diacronica è stata ribadita tra gli altri da Greenberg stesso (es. Greenberg 1995). È tuttavia soltanto negli ultimi decenni che l'approccio diacronico ha iniziato a essere più sistematicamente integrato con la pratica tipologica, soprattutto grazie a un sempre più fruttuoso dialogo con la linguistica storica e gli studi di grammaticalizzazione (v. Narrog 2017, Heine & Narrog 2018, Bisang *et al.* 2020).

Gli studi di grammaticalizzazione hanno dimostrato come in genere le strutture grammaticali si sviluppano a partire da una serie ricorrente di costruzioni di partenza e tramite processi analoghi in lingue diverse (es. Narrog 2017, Kuteva *et al.* 2019). Per esempio, la varietà di costruzioni dedicate alla codifica del possesso predicativo nelle lingue del mondo può essere ricondotta a 8 diversi schemi di origine (Heine 1997). Questi schemi in parte spiegano certe proprietà delle costruzioni di arrivo: il fatto che alcune lingue codificano il possessore allo stesso modo del locativo è diacronicamente dovuto al fatto che una costruzione di partenza 'Y si trova da X' viene reinterpretata come 'X ha Y' (Heine 1997: 91). In questo senso, lo studio delle costruzioni di partenza rende in parte conto del perché sincronicamente la variazione concerne un certo tipo di strutture e non altre.

La grammaticalizzazione (e più in genere la teoria del mutamento semantico in ambito lessicale) fornisce anche un importante strumento per spiegare perché nelle lingue del mondo le costruzioni grammaticali mostrano specifici ricorrenti pattern di polifunzionalità e non altri. Per esempio, sincronicamente è stato più volte osservato come costruzioni che codificano la funzione riflessiva frequentemente codificano anche la funzione reciproca e quella antipassiva. Ciò non è casuale ma deriva dal fatto che esiste un percorso diacronico ben preciso che porta le costruzioni riflessive a svilupparsi in costruzioni reciproche ed successivamente antipassive (Sansò 2017).

Studi diacronici di questo tipo hanno portato a ripensare più genericamente il ruolo dei principi sincronico-funzionali (§. 2.3) soggiacenti agli universali linguistici, alle gerarchie implicazionali e alle mappe semantiche (su queste ultime v. in particolare Luraghi 2014). Questo approccio alla tipologia, che possiamo definire *source-oriented*, si fonda sulla tesi che le regolarità nella distribuzione di certe strutture linguistiche non vada (esclusivamente) ricercata in principi sincronici funzionali, ma sia piuttosto dovuta a particolari vincoli nei

tipi di sorgente e di processi diacronici che portano all'emergere delle strutture in questione (cfr. Blevins 2004; Moravcsik 2010; Cristofaro 2010, 2013, 2017, 2019; Cristofaro & Zúñiga 2018; Mithun 2018; si veda anche Haspelmath 2019).

La tipologia diacronica *source-oriented* rappresenta quindi un importante punto di svolta rispetto al modello teorico della tipologia più tradizionale, perché considerazioni diacroniche diventano essenziali nello spiegare la struttura e i limiti della variazione interlinguistica.

4.2 Varietà sociolinguistiche e pattern di mutamento

Lo studio di fatti di variazione sociolinguistica è stato recentemente posto in relazione a processi diacronici. In particolare, in alcuni casi si osserva come le varianti di una certa variabile analizzata in diverse varietà di lingua rappresentino in realtà fasi diverse di una medesima costruzione nel suo processo di grammaticalizzazione (per una riflessione v. Nevalainen & Palander-Collin 2011; Poplack 2011). All'argomento è dedicato il contributo di Cerruti (in questo volume), a cui si rimanda.

A questa prospettiva si lega anche a quanto accennato nel § 2.1 in relazione a un possibile "rovesciamento" della variabile sociolinguistica tradizionalmente intesa. Così facendo, non si indagherebbe come una certa funzione si realizzi attraverso diverse forme ma, al contrario, come un'unica forma possa essere portatrice di diverse funzioni che correlano, al pari delle varianti tradizionali, con fatti extralinguistici.

In ambito italiano, l'argomento è stato affrontato da Cerruti (2011), che discute la presenza, accanto al modello della variabile sociolinguistica tradizionalmente intesa, di un'altra variabile che può dare conto di fatti di variazione sociolinguistica. Nel suo contributo, Cerruti (2011: 219) scrive infatti che:

Si potrebbero riconoscere due tipi distinti di variabile identificativi ciascuno di un diverso modo di variazione: un primo tipo – conforme alla concezione consueta di variabile sociolinguistica – in cui la variabile è intesa come ogni insieme di parole (le varianti) che designano uno stesso significato codificato nel sistema; un secondo tipo in cui la variabile è concepita come ogni insieme di significati, o meglio di sensi (le varianti), designati da una stessa parola (...). Una variabile del primo tipo, che potrebbe dirsi di tipo onomasiologico,

avrà dunque varianti formali; una variabile del secondo, di tipo semasiologico, varianti semantiche.

Cerruti propone questa distinzione in relazione alla dimensione lessicale. Essa, tuttavia, potrebbe essere declinata a elementi che veicolano significati discorsivi e pragmatici. Questo processo è tutt'altro che indolore: esso comporta che sia il ricercatore a determinare quali siano le funzioni associate alla forma. In altre parole, è il ricercatore a "ritagliare" da un certo spazio semantico varianti in termini di funzioni pragmatico-discorsive. Il problema era ben presente se si considera che già Weiner & Labov (1983:31) scrivevano che "theoretically, it should be possible to draw equal profit from cases where a single form is used with several meanings. But the possibility of accurate measurement is less immediate with semantic variation". Cionondimeno, l'adozione di questa prospettiva permette di fare a meno dell'equivalenza semantica e di individuare in maniera univoca i contesti da considerare determinati dall'occorrenza della forma oggetto di analisi.

Questa prospettiva, certamente debitrice degli studi di linguistica storica in cui si considera l'evoluzione nel tempo di un certo elemento lessicale, è stata declinata efficacemente a studi di stampo sociolinguistico.

Per l'ambito italiano (e romanzo), citiamo, tra gli altri, Moretti (2004) sul ciclo di *avere* dal latino ad alcune varietà romanze, Cerruti (2007) per alcune perifrasi aspettuali nell'italiano regionale piemontese, Poplack *et al.* (2018) sul congiuntivo in area romanza e Ballarè (2020) su alcuni usi di *mica* in varietà diatopiche diverse.

5. *Un quadro d'insieme*

L'esistenza di affinità teoriche e metodologiche tra la sociolinguistica e la tipologia ha portato in tempi più recenti a un sempre proficuo scambio tra le due discipline. Negli ultimi anni, diversi contributi hanno discusso in maniera sistematica i possibili benefici che le due discipline possono trarre da un'interazione più sistematica (v. ad es. i lavori presentati in Kortmann 2004a e la riflessione portata avanti in Trudgill 2011), e diversi progetti sono nati ispirati a queste linee di ricerca (per cui v. anche Grandi e Di Garbo *et al.* in questo volume).

Nel resto di questa sezione, offriamo una breve panoramica di alcuni sviluppi recenti nel rapporto tra la sociolinguistica e la tipologia.

5.1 Dalla tipologia alla sociolinguistica

La progressiva apertura della sociolinguistica (e, più in generale, l'osservazione di fatti di variazione interni ad un'unica lingua) a considerazioni tipologiche ha il vantaggio di introdurre nello studio delle variabili dei *tertia comparationis* indipendenti dalla descrizione di singole lingue o famiglie linguistiche. Già Bossong (1991: 143) osservava che “broad typological comparison of data from many genetically and structurally different languages is necessary in order to be able to describe phenomena of single languages as what they really are”.

È d'obbligo citare il volume curato da Kortmann (2004a) che ha come obiettivo programmatico l'invito ad un confronto sistematico tra chi si interessa di variazione intra- e interlinguistica. Sin dalle prime righe del contributo introduttivo, infatti, si scrive che “the purpose of this invitation is to bring together for the first time two research traditions in the study of language variation (and change) which so far have largely worked independently of each other, to make them enrich and provide new vistas for each other” (Kortmann 2004b: 1). Tra i temi affrontati nel volume, merita qualche riga quello relativo ai *vernacular universals* (Chambers 2004, 2009) che sono da considerarsi come tratti linguistici che caratterizzerebbero varietà substandard di lingue diverse. Secondo Chambers, questi universali comprendono “a small number of phonological and grammatical processes [that] recur in vernaculars wherever they are spoken. This conclusion follows from the observation that, no matter where in the world the vernaculars are spoken [...] these features inevitably occur” (Chambers 2004: 128). Alcuni di questi universali sarebbero, ad esempio, la regolarizzazione di coniugazioni, il singolare di default (come ad es. in *they was the last ones*), la negazione multipla e la cancellazione della copula (Chambers 2004: 129). In letteratura, ci si è concentrati principalmente sulla presenza di *vernacular universals* in varietà di inglese (v. ad es. Szmrecsanyi & Kortmann 2009), anche se recentemente sono stati discussi anche in relazione ad alcune strutture dell'italiano popolare (Ballarè & Gorla 2019) e alle frasi relative e alla gerarchia di accessibilità in diverse lingue europee (Ballarè & Larrivé in stampa).

In ambito italiano, a partire dagli anni Ottanta, gli strumenti della tipologia sono stati adottati per trattare fatti di variazione intralinguistica. Singole costruzioni sono state analizzate con gli strumenti della tipologia linguistica (v. ad es. la discussione sull'accusativo preposizionale di Berretta 1989) e costruzioni in competizione sono state considerate (anche) in relazione a caratteristiche tipologiche dell'italiano (v. ad es. Bernini 1992 su alcune strategie di negazione non canonica). Tratti strutturali dell'italiano neo-standard sono stati discussi in chiave tipologica (Berretta 1994) e studi di chiara impostazione sociolinguistica hanno posto l'attenzione su fatti di rilievo anche per la variazione interlinguistica – si vedano, tra gli altri, i lavori di Berruto (1983, 1990) su varietà substandard e semplificazione linguistica. Più recentemente, questa prospettiva è stata adottata, ad esempio, nello studio delle frasi relative dell'italiano, che vengono realizzate con diverse configurazioni tipologiche in diverse varietà di lingua (Cerruti 2017). A conclusione, possiamo citare il progetto PRIN 2017 UniverS-ITA che ha come obiettivo programmatico lo studio della lingua degli studenti universitari, integrando la prospettiva tipologica e quella sociolinguistica.

5.2 Dalla sociolinguistica alla tipologia

Due sono gli ambiti principali in cui si evidenzia un apporto benefico della sociolinguistica alla ricerca tipologica.

In primo luogo, numerosi studi, soprattutto a partire da Trudgill (2011) hanno messo in evidenza come lo studio della diversità interlinguistica non può prescindere da considerazioni sociali, e che particolari tratti linguistici possano essere in parte spiegati con riferimento a fattori non linguistici di natura prettamente storico-geografico o sociale. In questa dimensione si collocano ad esempio studi che cercano una correlazione tra particolari condizioni climatiche e il numero di parlanti non nativi in una comunità con la dimensione dell'inventario fonologico di una lingua (es. Wray & George 2007; Everett *et al.* 2015; v. Di Garbo *et al.* nel presente volume per una discussione dettagliata di questi temi). Come osserva Grandi (questo volume), considerazioni di questo genere hanno anche importanti ripercussioni metodologiche, per esempio sui criteri di costruzione dei campioni tipologici. Più in generale, è a lungo stato dimostrato come considerazioni di tipo geografico, o meglio areale, debbano essere tenute in

conto nello studio della distribuzione interlinguistica ti particolari tratti (es. Nichols 1992; v. per una rassegna Koptjeskaja-Tamm 2010; Bickel 2017).

Una seconda linea di ricerca, anch'essa con importanti conseguenze metodologiche sulla prassi di campionamento, riguarda l'integrazione della nozione di varietà di lingua negli studi tipologici. Per la tipologia, la lezione della sociolinguistica consiste nell'opportunità di tenere conto, per le lingue con una situazione sociolinguistica complessa, (anche) di varietà non standard, le quali possono offrire una gamma di fenomeni più ampia rispetto a quella presentata dalla varietà standard codificata nelle grammatiche e possono portare a ricalibrare valutazioni di carattere tipologico basate su varietà standardizzate (v. già Bossong 1985 e più recentemente Murelli & Kortmann 2011; Seiler 2019; Grandi in questo volume). È noto, infatti, come il confronto tra varietà non standard, tipicamente orali, metta in luce più facilmente caratteristiche comuni condivise tra lingue anche tipologicamente distanti (v. ad es. Auer 1990; Auer & Maschler 2013 per un confronto tra ebraico moderno e tedesco; Ballarè & Inglese 2021 sulle relative locative).

6. *Note conclusive*

In questo contributo abbiamo messo in luce alcuni aspetti che avvicinano la sociolinguistica e la tipologia. Sebbene le due discipline abbiano premesse, scopi, e ambiti di indagine tradizionalmente ben differenziati, in realtà come abbiamo avuto modo di discutere nel corso di questo lavoro, esse presentano numerose affinità. In prima istanza, abbiamo mostrato come entrambe le discipline hanno elaborato dei modelli teorici per lo studio della variazione, mettendo l'accento su problemi comuni – primo fra tutti, la delimitazione dei confini dell'indagine – e le strategie elaborate per superarle. Secondariamente, sono state presentate le risorse tipicamente usate dalle due discipline per l'analisi della variazione, ovvero atlanti, questionari e corpora. Nel quarto paragrafo, è stata introdotta la dimensione diacronica e si è discusso di come essa sia stata efficacemente posta in relazione con quella sincronica nel gettare luce sui fenomeni di variazione, sia interlinguistica che intralinguistica. Infine, abbiamo presentato una breve

rassegna di lavori di sociolinguistica e tipologia che hanno beneficiato dell'apporto dell'altra disciplina.

In conclusione, auspichiamo che queste pagine (nonché gli altri contributi presenti in questo volume) possano fungere da spunto di riflessione per un approccio allo studio della variazione che miri sempre più a integrare prospettive diverse.

Ringraziamenti

Una versione preliminare di questo capitolo è stata discussa con Massimo Cerruti, Nicola Grandi e Simone Mattioli, a cui va il nostro ringraziamento. Errori e imprecisioni sono ovviamente da attribuirsi solamente a chi scrive.

Riferimenti bibliografici

- Alfieri, Luca. 2014. Qualifying modifier encoding and adjectival typology. In Simone, Raffaele & Masini, Francesca (a cura di), *Word Classes: Nature, typology and representations*, 119-138. Amsterdam: Benjamins.
- Alfieri, Luca, Arcodia, Giorgio Francesco & Ramat, Paolo (a cura di). 2021. *Linguistic Categories, Language Description and Linguistic Typology*. Amsterdam: Benjamins.
- Arkadiev, Peter & Klamer, Marian. 2018. Morphological theory and typology. In Audring, Jenny & Masini, Francesca (a cura di), *The Oxford Handbook of Morphological Theory*, 436-454. Oxford: Oxford University Press.
- Auer, Peter. 1990. "Einige umgangssprachliche Phänomene der türkischen Syntax und Möglichkeiten ihrer Erklärung aus 'natürlichen' Prinzipien". In Boretzky, Norbert & Enninger, Werner & Stolz, Thomas (a cura di), *Spielarten der Natürlichkeit-Spielarten der Ökonomie. Beiträge zum 5. Essener Kolloquium*, vol. 2, 271-298. Bochum: Brockmeyer.
- Auer, Peter & Maschler, Yael. 2013. Discourse or grammar? VS patterns in spoken Hebrew and spoken German narratives. *Language Science* 37. 147-181.
- Ballarè, Silvia. 2020. *Esiti del ciclo di Jespersen in area italo-romanza: grammaticalizzazione e contatto linguistico*. Bologna – Cesena: Caissa Italia.

- Ballarè, Silvia & Gorla, Eugenio. 2019. *C'era i fascisti e i tedeschi*: instances of linguistic simplification in a corpus of *Italiano popolare*. In Villena Ponsoda, Juan Andrés & Díaz-Montesinos, Francisco & Ávila-Muñoz, Antonio-Manuel & Vida-Castro, Matilde (a cura di), *Language Variation – European perspectives VII. Studies in Language Variation Series*, 73-86. Amsterdam: Benjamins.
- Ballarè, Silvia & Inglese, Guglielmo. 2021. The development of locative relative markers. From typology to sociolinguistics (and back). *Studies in Language*. <https://doi.org/10.1075/sl.20013.bal>
- Ballarè, Silvia & Larrivé, Pierre. (in stampa). Register impacts syntax: Scaling the accessibility hierarchy of relatives. *Italian Journal of Linguistics*.
- Barbiers, Sjef & Bennis, Hans & de Vogelaer, Gunther & Devos, Magda & van der Ham, Margreet. 2005. *Syntactische Atlas van de Nederlandse Dialecten*, Deel I. Amsterdam: Amsterdam University Press.
- Bartoli, Matteo G. (dir.). 1995. *Atlante linguistico italiano*, 7 voll., a cura di Pellis, Ugo & Massobio, Lorenzo. Roma: Istituto Poligrafico e Zecca dello Stato.
- Berlin, Brent & Kay, Paul. 1969. *Basic color terms: their universality and evolution*. Berkeley, CA: University of California Press.
- Bernini, Giuliano. 1992. Forme concorrenti di negazione in italiano. In Moretti, Bruno & Petrini, Dario & Bianconi, Sandro (a cura di), *Linee di tendenza dell'italiano contemporaneo. Atti del XXV Congresso Internazionale di Studi della SLI, Lugano 19.21 settembre 1991*, 191-215. Roma: Bulzoni.
- Berretta, Monica. 1989. Sulla presenza dell'accusativo preposizionale in italiano. *Vox romanica* 48. 13-37.
- Berretta, Monica. 1994. Correlazioni tipologiche fra tratti morfosintattici dell'italiano 'neo-standard'. In Holtus, Günter & Radtke, Edgar (a cura di), *Sprachprognostik und das 'italiano di domani'. Prospettive per una linguistica 'prognostica'*, 125-152. Tübingen: Niemeyer.
- Berruto, Gaetano. 1983. L'italiano popolare e la semplificazione linguistica. *Vox romanica* 42. 38-79.
- Berruto, Gaetano. 1990. Semplificazione linguistica e varietà sub-standard. In Holtus, Günter & Radtke, Edgar (a cura di), *Sprachlicher Substandard III. Standard, Substandard und Varietätenlinguistik*, 17-43. Tübingen: Niemeyer.
- Berruto, Gaetano. 2007 [1995]. *Fondamenti di sociolinguistica*. Bari: Laterza.
- Berruto, Gaetano. 2009. Περὶ συντάξεως. Sintassi e variazione. In Ferrari, Angela (a cura di), *Sintassi storica e sincronica dell'italiano*.

- Subordinazione, coordinazione, giustapposizione, Atti del X Congresso della Società Internazionale di Linguistica e Filologia italiana (Basilea, 30 giugno-3 luglio 2008)*, 21-58. Firenze: Cesati.
- Berruto, Gaetano. 2010. Identifying dimensions of linguistic variation in a language space. In Auer, Peter & Schmidt, Jürgen Erich (a cura di), *Language and Space, Part I: Theories and Methods. An International Handbook of Linguistic Variation*, 226-241. Berlin - New York: de Gruyter.
- Berruto, Gaetano. 2012. *Sociolinguistica dell'italiano contemporaneo. Nuova edizione*. Roma: Carocci.
- Berruto, Gaetano & Cerruti, Massimo. 2015. *Manuale di sociolinguistica*. Torino: UTET.
- Bickel, Balthasar. 2010. Grammatical relations typology. In Song, Jae Jung (a cura di), *The Oxford Handbook of Linguistic Typology*, 399-444. Oxford: Oxford University Press.
- Bickel, Balthasar. 2017. Areas and universals. In Hickey, Raymond (a cura di), *The Cambridge Handbook of Areal Linguistics*, 40-55. Cambridge: Cambridge University Press.
- Bisang, Walter, Malchukov, Andrej & the Mainz Grammaticalization Project team (Iris Rieder, Linlin Sun, Marvin Martiny, Svenja Luell). 2020. Position paper: Universal and areal patterns in grammaticalization. In Bisang, Walter & Malchuko, Andrej (a cura di), *Grammaticalization Scenarios: Cross-linguistic Variation and Universal Tendencies*, 1-88. Berlin - New York: de Gruyter.
- Blevins, Juliette. 2004. *Evolutionary phonology: the emergence of sound patterns*. Cambridge: Cambridge University Press.
- Bohnemeyer, Jürgen & Caelen, Martijn. 1999. The ECOM clips: A stimulus for the linguistic coding of event complexity. In Wilkins, David (a cura di), *Manual for the 1999 Field Season*, 74-86. Nijmegen: Max Planck Institute for Psycholinguistics.
- Bossong, Georg. 1985. *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.
- Bossong, Georg. 1991. Differential Object Marking in Romance and Beyond. In Kibbee, Douglas & Dieter Wanner (a cura di), *New Analyses in Romance Linguistics*, 143-170. Amsterdam: Benjamins.
- Cameron, Richard & Schwenker, Scott. 2013. Pragmatics and variationist sociolinguistics. Bayley, Robert & Cameron, Richard & Lucas, Celi (a cura di), *The Oxford Handbook of Sociolinguistics*, 464-483. Oxford: Oxford University Press.

- Carling, Gerd (a cura di) 2017. *Diachronic Atlas of Comparative Linguistics Online*. Lund: Lund University. (<https://diac.ht.lu.se/>) (Consultato il 7.06.2021).
- Cerruti, Massimo. 2007. Sulla caratterizzazione aspettuale e la variabilità sociale d'uso di alcune perifrasi diatopicamente marcate. *Archivio Glottologico Italiano* 92(2) 203-247.
- Cerruti, Massimo. 2009. *Strutture dell'italiano regionale. Morfosintassi di una varietà diatopica in prospettiva sociolinguistica*. Frankfurt am Main: Lang.
- Cerruti, Massimo. 2010. Teoria dei prototipi e variazione linguistica: la categoria di scala di implicazione in prospettiva tipologica. *Vox Romanica* 69. 25-46.
- Cerruti, Massimo. 2011. Il concetto di variabile sociolinguistica a livello del lessico. *Studi italiani di linguistica teorica e applicata* 40(2). 211-231.
- Cerruti, Massimo. 2017. Changes from below, changes from above: relative constructions in contemporary Italian. In Cerruti, Massimo & Crocco, Claudia & Marzo, Stefania (a cura di), *Towards a New Standard*, 62-88. Berlin-New York: de Gruyter.
- Cerruti, Massimo & Vietti, Alessandro. (in stampa). Identifying language varieties: Coexisting standards in spoken Italian. In Beaman, Karen & Guy, Gregory R. (a cura di), *The coherence of linguistic communities: orderly heterogeneity and social meaning*. London: Routledge.
- Chambers, Jack K. 2003. *Sociolinguistic theory: linguistic variation and its social significance*. Malden-Oxford: Blackwell Publishers.
- Chambers, Jack. 2004. Dynamic typology and vernacular universals. In Kortmann, Bernd (a cura di), *Dialectology Meets Typology: Dialect Grammar from a Cross-Linguistic Perspective*, 128-145. Berlin-New York: de Gruyter.
- Chambers, Jack. 2009. Cognition and the Linguistic Continuum from Vernacular to Standard. In Filppula, Marrku & Klemola, Juhani & Paulasto, Heli. *Vernacular Universals and Language Contacts*, 19-32. London-New York: Routledge.
- Chambers, Jack & Trudgill, Peter. 2004 [1998]. *Dialectology*. Cambridge: Cambridge University Press.
- Comrie, Bernard & Kuteva, Tania. 2013. Relativization Strategies. In Dryer, Matthew S. & Haspelmath, Martin (a cura di), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute (<http://wals.info/chapter/s8>) (Consultato il 10.04.2021).
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

- Coseriu, Eugenio. 1973. Sulla tipologia linguistica di Wilhelm von Humboldt: contributo alla critica della tradizione linguistica. *Lingua e stile* 8. 235-266.
- Cristofaro, Sonia & Zúñiga, Fernando. 2018. Synchronic vs. diachronic approaches to typological hierarchies. In Cristofaro, Sonia & Zúñiga, Fernando (a cura di), *Typological hierarchies in synchrony and diachrony*, 4-27. Amsterdam: Benjamins
- Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.
- Cristofaro, Sonia. 2010. Semantic maps and mental representation. *Linguistic Discovery* 8(1). Doi: 10.1349/PS1.1537-0852.A.345
- Cristofaro, Sonia. 2013. The referential hierarchy: Reviewing the evidence in diachronic perspective. In Bakker, Dik & Haspelmath, Martin (a cura di), *Languages Across Boundaries: Studies in the Memory of Anna Siewierska*, 69-93. Berlin - New York: de Gruyter.
- Cristofaro, Sonia. 2017. Implicational universals and dependencies. In Enfield, Nick J. (a cura di), *Dependencies in language: On the causal ontology of linguistic systems*, 9-22. Berlin: Language Science Press.
- Cristofaro, Sonia. 2019. Taking diachronic evidence seriously: Result-oriented vs. source-oriented explanations of typological universals. In Schmidtke-Bode, Karsten, Levshina, Natalia, Michaelis, Susanne Maria, & Seržant, Ilja A. (a cura di), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 25-46. Berlin: Language Science Press.
- Croft, William. 2000. *Explaining language change: an evolutionary approach*. Harlow: Pearson Longman.
- Croft, William. 2003. *Typology and universals*. Cambridge: Cambridge University Press.
- Croft, William. 2016. Comparative concepts and language-specific categories: Theory and practice. *Linguistic Typology* 20(2). 377-393.
- Cysouw, Michael & Wälchli, Bernhard. 2007. Parallel texts: Using translational equivalents in linguistic typology. *Sprachtypologie und Universalienforschung (STUF)* 60(2). 95-99.
- Cysouw, Michael. 2007. Building semantic maps: The case of person marking. In Wälchli, Bernhard & Miestamo, Matti (a cura di), *New Challenges in Typology*, 225-248. Berlin - New York: de Gruyter.
- Cysouw, Michael. 2010. Semantic maps as metrics on meaning. *Linguistic Discovery* 8(1). Doi: 10.1349/PS1.1537-0852.A.346.
- Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Blackwell.

- Daniel, Michael. 2010. Linguistic Typology and the Study of Language. In Song, Jae Jung (a cura di), *The Oxford Handbook of Linguistic Typology*, 43-68. Oxford: Oxford University Press.
- Dixon, R.M.W. 1982. *Where Have All the Adjectives Gone?, and Other Essays in Semantics and Syntax*. Berlin: Mouton.
- Dryer, Matthew S. & Haspelmath, Martin (a cura di). 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology (wals.info).
- Evans, Nicholas. 2020. Introduction: Why the comparability problem is central in typology. *Linguistic Typology* 24(3). 417-425.
- Everett, Caleb, Blasi, Damián E. & Roberts, Seán G. 2015. Climate, vocal folds, and tonal languages: connecting the physiological and geographic dots. *PNAS* 112. 1322-1327.
- François, Alexnadre. 2008. Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. In Vanhove, Martine (a cura di), *From Polysemy to Semantic Change: Towards a typology of lexical semantics associations*, 163-215. Amsterdam: Benjamins.
- Georgakopoulos, Thanasis & Polis, Stéphane. 2018. The semantic map mode. State of the art and future avenues for linguistic research. *Language and Linguistics Compass* 12(2). e12270.
- Gerdes, Kim, Kahane, Sylvain & Chen, Xinying. 2021. Typometrics: From Implicational to Quantitative Universals in Word Order Typology. *Glossa: a journal of general linguistics* 6(1). 17.
- Ghyselen, Anne-Sophie & De Vogelaer, Gunther. 2018. Seeking systematicity in variation: theoretical and methodological considerations on the “variety” concept. *Frontier in psychology* 9. 385.
- Giacalone Ramat, Anna, Mauri, Caterina & Molinelli, Piera (a cura di). 2012. *Synchrony and Diachrony. A dynamic interface*. Amsterdam: Benjamins.
- Gil, David. 2016. Describing languoids: When incommensurability meets the language-dialect continuum. *Linguistic Typology* 20(2). 439-462.
- Goria, Eugenio & Mauri, Caterina. 2018. Il corpus KIParla: una nuova risorsa per lo studio dell'italiano parlato. In Masini, Francesca & Tamburini, Fabio (a cura di), *CLUB Working Papers in Linguistics* vol. 2, 96-110. Bologna: CLUB – Circolo Linguistico dell'Università di Bologna.
- Grandi, Nicola. 2014. *Fondamenti di tipologia linguistica*. Roma: Carocci.
- Grandi, Nicola. 2018. Su alcune possibili interazioni tra tipologia e sociolinguistica. In Gianollo, Chiara & Mauri, Caterina (a cura di),

- CLUB Working Papers in Linguistics*, 257-265. Alma Mater Studiorum – Univesità di Bologna.
- Grandi, Nicola. 2020. La diversità inevitabile. La variazione linguistica tra tipologia e sociolinguistica. *Italiano LinguaDue* 12. 416-429.
- Greenberg, Joseph H. 1960. A Quantitative Approach to the Morphological Typology of Language. *International Journal of American Linguistics* 26(3). 178-194.
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, Joseph H. (a cura di), *Universals of language*, 73-113. Cambridge: MIT Press.
- Greenberg, Joseph H. 1995. The diachronic typological approach to language. In Shibatani, Masayoshi & Bynon, Theodora (a cura di), *Approaches to language typology*, 145-166. Oxford: Clarendon Press.
- Haspelmath, Martin. 1993. More on the typology of the inchoative/causative verb alternations. In Comrie, Bernard & Polinsky, Maria (a cura di), *Causatives and Transitivity*, 87-120. Amsterdam: Benjamins.
- Haspelmath, Martin. 2003. The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. In Michael Tomasello (a cura di), *The new psychology of language: Cognitive and Functional Approaches to Language Structure*, Vol. 2, 211-242. Mahwah, NJ: Lawrence Erlbaum.
- Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42(1). 25-70.
- Haspelmath, Martin. 2008. A frequentist explanation of some universals of reflexive marking. *Linguistic Discovery* 6(1). 40-63.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663-687.
- Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31-80.
- Haspelmath, Martin (dir). 2014-. *Studies in Diversity Linguistics*. Berlin: Language Science Press.
- Haspelmath, Martin. 2016. Universals of causative and anticausative verb formation and the spontaneity scale. *Lingua Posnaniensis* 58 (2). 33-63.
- Haspelmath, Martin. 2018. How comparative concepts and descriptive linguistic categories are different. In Van Olmen, Daniël, Mortelmans, Tanja & Brisard, Frank (a cura di), *Aspects of linguistic variation*, 83-114. Berlin - New York: de Gruyter.
- Haspelmath, Martin. 2019. Can cross-linguistic regularities be explained by change constraints? In Schmidtke-Bode, Karsten, Levshina, Natalia,

- Michaelis, Susanne Maria, & Seržant, Ilja A. (a cura di), *Explanation in typology: Diachronic sources, functional motivations and the nature of the evidence*, 1-23. Berlin: Language Science Press.
- Haspelmath, Martin. 2021. Explaining grammatical coding asymmetries: Form-frequency correspondencies and predictability. *Journal of Linguistics*. doi: 10.1017/S0022226720000535
- Haspelmath, Martin, Dryer, Matthew S., Gil, David & Comrie, Bernard (a cura di). 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- Hinskens, Frans & Guy, Gregory R. (a cura di). 2016. *Coherence, covariation and bricolage. Various approaches to the systematicity of language variation*. Numeri speciali di *Lingua* 172-173.
- Heine, Bernd. 1997. *Cognitive foundations of grammar*. Oxford: Oxford University Press.
- Heine, Bernd & Narrog, Heiko (a cura di). 2018. *The Oxford Handbook of Grammaticalization*. Oxford: Oxford University Press.
- Hopper, Paul J. & Traugott, Elizabeth Closs. 2003. *Grammaticalization. Second edition*. Cambridge: Cambridge University Press.
- Hudson, Richard A. 1998 [1980]. *Sociolinguistica*. Bologna: il Mulino (trad. di *Sociolinguistics*. 1980. Cambridge: Cambridge University Press).
- Iannàcaro, Gabriele. 2000. Per una semantica più puntuale del concetto di 'dato linguistico'. *Quanderni di semantica XXI*(1). 51-79.
- Jaberg, Karl & Jud, Jakob. 1928-1940. *Sprach- und Sachtlas Italiens und der Südschweiz*, 8 voll. Zofingen: Ringier.
- Keenan, Edward & Comrie, Bernard. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8. 63-99.
- Keenan, Edward & Comrie, Bernard. 1979. Noun phrase accessibility revisited. *Language* 55(3). 649-664.
- Koptjevskaja-Tamm, Maria 2012. New directions in lexical typology. *Linguistics* 50(3). 373 – 394.
- Koptjeskaja-Tamm, Maria, Rakhilina, Ekaterina & Vanhove, Martine. 2015. The semantics of lexical typology. In Nick Riemer (a cura di), *The Routledge Handbook of Semantics*, 434-454. London: Routledge.
- Kortmann, Bernd & Lunkenheimer, Krestin & Ehret, Katharina (a cura di). 2020. *The Electronic World Atlas of Varieties of English* (<https://ewave-atlas.org/>).

- Kortmann, Bernd (a cura di). 2004a. *Dialectology Meets Typology: Dialect Grammar from a Cross-Linguistic Perspective*. Berlin-New York: de Gruyter.
- Kortmann, Bernd. 2004b. *Introduction*. In Kortmann, Bernd (a cura di), *Dialectology Meets Typology: Dialect Grammar from a Cross-Linguistic Perspective*. Berlin - New York: de Gruyter.
- Kortmann, Bernd & Lunkenheimer, Kerstin & Ehret, Katharina (a cura di). 2020. *The Electronic World Atlas of Varieties of English* (ewave-atlas.org).
- Kuteva, Tania, Heine, Bernd, Hong, Bo, Long, Haiping, Narrog, Heiko & Rhee, Seongha. 2019. *World lexicon of grammaticalization. Second, extensively revised and updated edition*. Cambridge: Cambridge University Press.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Washington: Center for Applied Linguistics.
- Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William. 1978. Where does the sociolinguistic variable stop? A response to Beatriz Lavandera. *Working paper in sociolinguistics* 44. 6-21.
- Labov, William. 1982. Building on empirical foundations. In Lehmann, Winfried P. & Malkiel, Yakov (a cura di), *Perspective on historical linguistics*, 17-92. Amsterdam: Benjamins.
- Labov, William. 2004. Quantitative reasoning in linguistics. In Ammon, Ulrich & Dittmar, Norbert & Mattheier, Klaus J. (a cura di), *Sociolinguistics/Soziolinguistik: An International Handbook of the Science of Language and Society*. Volume 1, 6-22. Berlin - New York: de Gruyter.
- Labov, William & Ash, Sharon & Boberg, Charles. 2006. *The Atlas of North American English: Phonetics, phonology and sound change*. Berlin - New York: de Gruyter.
- Lameli, Alfred. 2010. Linguistic Atlases – traditional and modern. In Auer, Peter & Schmidt, Jürgen Erich (a cura di), *Language and Space, Part I: Theories and Methods. An International Handbook of Linguistic Variation*, 567-592. Berlin - New York: de Gruyter.
- Lameli, Alfred & Kehrein, Roland & Rabanus, Stefan (a cura di). 2011. *Language and Space, Part II: Language mapping*. Berlin - New York: de Gruyter.
- Lavandera, Beatriz. 1978. Where does the sociolinguistic variable stop?. *Language in Society* 7(2). 171-182.

- Lehmann, Christian. 2015 [1995]. *Thoughts on grammaticalization*, terza edizione. Berlin: Language Science Press.
- Levshina, Natalia. 2019. Token-based typology and word order entropy: A study based on Universal Dependencies. *Linguistic Typology* 23(3). 533-572.
- Levshina, Natalia. 2020. Corpus-based typology: applications, challenges and some solutions. *Linguistic Typology*. <https://doi.org/10.1515/lingty-2020-0118>.
- List, Johann-Mattis, Greenhill, Simon J., Anderson, Cormac, Mayer, Thomas, Tresoldi, Tiago & Forkel, Robert. 2018. CLICS2: An improved database of cross-linguistic colexifications assembling lexical data with the help of cross-linguistic data formats. *Linguistic Typology* 22(2). 277-330
- Luraghi, Silvia. 2010. Causes of language change. In Luraghi, Silvia & Bubenik, Vit (a cura di), *Continuum Companion to Historical Linguistics*, 354-366. London - New York: Continuum.
- Luraghi, Silvia. 2014. Plotting diachronic semantic maps: the role of metaphors. In Luraghi, Silvia & Narrog, Heiko (a cura di), *Perspectives on Semantic Roles*, 99-150. Amsterdam: Benjamins.
- Malchukov, Andrej. 2005. Case Pattern Splits, Verb Types and Construction Competition. In Amberber, Mengistu & de Hoop, Helen (a cura di), *Competition and variation in natural languages: The case for case*, 73-117. Amsterdam: Elsevier.
- Mauri, Caterina & Ballarè, Silvia & Gorla, Eugenio & Cerruti, Massimo & Suriano, Francesco. 2019. KIParla corpus: a new resource for spoken Italian. In Bernardi, Raffaella & Navigli, Roberto & Semeraro, Giovanni (a cura di), *Proceedings of the Sixth Italian Conference on Computational Linguistics*. DOI: <http://ceur-ws.org/Vol-2481/paper45.pdf>.
- Meyerhoff, Miriam & Schlee, Erik & MacKenzie. 2015. *Doing sociolinguistics. A practical guide to data collection and analysis*. London: Routledge.
- Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.) 2013. *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<http://apics-online.info>) (Consultato 31.05.2021).
- Miestamo, Matti. 2005. *Standard negation: the negation of declarative verbal main clauses in a typological perspective*. Berlin - New York: de Gruyter.
- Mithun, Marianne. 2018. Deconstructing teleology: The place of synchronic usage patterns among processes of diachronic development. In Cristofaro, Sonia & Zúñiga, Fernando (a cura di), *Typological hierarchies in synchrony and diachrony*, 111-128. Amsterdam: Benjamins.

- Moran, Steven & McCloy, Daniel (a cura di) 2019. PHOIBLE 2.0. Jena: Max Planck Institute for the Science of Human History. (<http://phoible.org>) (Consultato il 07.06.2021)
- Moravcsik, Edith A. 2010. Explaining Language Universals. In Song, Jae Jung (a cura di), *The Oxford Handbook of Linguistic Typology*, 69-89. Oxford: Oxford University Press
- Moravcsik, Edith A. 2016. On linguistic categories. *Linguistic Typology* 20(2). 417-425
- Moretti, Bruno. 2004. Il 'ciclo di avere'. Costanti e variazioni dal latino all'italiano moderno. *Rivista italiana di linguistica e dialettologia* VI. 141-160.
- Murelli, Adriano & Kortmann, Bernd. 2011. Non-standard varieties in the areal typology of Europe. In Kortmann, Bernd & van der Auwera, Joan (a cura di), *The Languages and Linguistics of Europe. A Comprehensive Guide*, 525-544. Berlin-New York: de Gruyter.
- Muysken, Pieter & Hammarström, Harald & Krasnoukhova, Olga & Müller, Neele & Brichall, Joshua & van de Kerke, Simon & O'Connor, Loretta & Danielsen, Swintha & van Gijn, Rik & Saad, George. 2016. *South American Indigenous Language Structures (SAILS) Online*. Jena: Max Planck Institute for the Science of Human History. (<http://sails.cld.org>) (Consultato il 07.06.2021).
- Narrog, Heiko. 2017. Typology and Grammaticalization. In Aikhenvald, Alexandra A. & Dixon, R.M.W. (a cura di), *The Cambridge Handbook of Linguistic Typology*, 151-177. Cambridge: Cambridge University Press.
- Nevalainen, Terttu & Palander-Collin, Minna. 2011. Grammaticalization and sociolinguistics. In Heine, Bernd & Narrog, Heiko (a cura di), *The Oxford Handbook of Grammaticalization*, 118-128. Oxford: Oxford University Press.
- Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.
- Paschen, Ludger, Delafontaine, François, Draxler, Christoph, Fuchs, Susanne, Stave, Matthew & Seifart, Frank. 2020. Building a time-aligned cross-linguistic reference corpus from language documentation data (DoReCo). *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*, 2657-2666. Marseille: European Language Resources Association. <http://www.lrec-conf.org/proceedings/lrec2020/pdf/2020.lrec-1.324.pdf>
- Pichler, Heike. 2013. *The structure of discourse-pragmatic variation*. Amsterdam: Benjamins.

- Poplack, Shana. 2011. Grammaticalization and linguistic variation. Grammaticalization and sociolinguistics. In Heine, Bernd & Narrog, Heiko (a cura di), *The Oxford Handbook of Grammaticalization*, 209-224. Oxford: Oxford University Press.
- Poplack, Shana. 2021. Le corpus comme portail pour l'étude de la variation (socio)linguistique. *Corpus [en ligne]* 22 DOI: <https://doi.org/10.4000/corpus.5422>.
- Poplack, Shana & Torres Cacoullos, Rena & Dion, Nathalie & de Andrade Berlinck, Rosane & Digesto, Salvatore & Lacasse, Dora & Steuck, Jonathan. 2018. Variation and Grammaticalization in Romance: A cross-linguistic study of the subjunctive. In Ayres-Bennet, Wendy & Carruther, Janice (a cura di), *Manual in Linguistics: Romance Sociolinguistics*, 217-252. Berlin-New York: de Gruyter.
- Ramat, Paolo. 2010. The (Early) History of Linguistic Typology. In Song, Jae Jung (a cura di), *The Oxford Handbook of Linguistic Typology*, 9-24. Oxford: Oxford University Press.
- Rickford, John R. 2002. Implicational Scale. In Chambers, Jack K. & Trudgill, Peter & Schilling-Estes, Natalie (a cura di), *The handbook of language variation and change*, 142-167. Oxford: Blackwell.
- Rießler, Michael. 2016. *Adjective attribution*. Berlin: Language Science Press.
- Round, Erich R. & Corbett, Greville G. 2020. Comparability and measurement in typological science: The bright future for linguistics. *Linguistic Typology* 24(3). 489-525.
- Sanga, Glauco. 1991. I metodi della ricerca sul campo. *Rivista Italiana di Dialettologia* 15. 165-181.
- Sankoff, Gillian. 1972. Above and beyond phonology in variable rules. In Bailey, Charles-James & Shuy, Roger W. (a cura di), *New ways of analyzing variation in English*, 44-61. Washington: Georgetown University Press.
- Sansò, Andrea. 2017. Where do antipassives constructions come from? A study in diachronic typology. *Diachronica* 34(2). 175-218.
- Sansò, Andrea. 2018. Explaining the diversity of antipassives: Formal grammar vs. (diachronic) typology. *Language and Linguistics Compass* 12(6). e12277.
- Sansò, Andrea. 2020. *I segnali discorsivi*. Roma: Carocci.
- Sapir, Edward. 1921. Language. *An Introduction to the Study of Speech*. New York: Harcourt, Brace & World.
- Saussure, Ferdinand de, 2016 [1916]. *Cours de linguistique générale*, Bally, Charles & Sechehaye, Albert (a cura di). Paris: Payot & Rivages.

- Schneider, Klaus P. & Barron, Anne (a cura di). 2008. *Variational pragmatics*. Amsterdam: Benjamins.
- Seiler, Guido. 2018. Synchrony and diachrony: two outdated dimensions? In Glessgen, Martin, Kabatek, Johannes, Völker, Harald (a cura di), *Repenser la variation linguistique. Actes du Colloque DIA IV à Zurich (12-14 sept. 2016)*, 77-96. Strasbourg: Éditions de linguistique et de philologie.
- Seiler, Guido. 2019. Non-Standard Average European. In Nievergelt, Andreas & Ludwig Rübekeil (a cura di), *'athe in palice, athe in anderu sumeuuelicheru stedi'. Raum und Sprache. Festschrift für Elvira Glaser zum 65 Geburtstag*, 541-554. Heidelberg: Winter.
- Siegel, Jeff, Szmrecsanyi, Benedikt & Kortmann, Bernd. 2014. Measuring analyticity and syntheticity in creoles. *Journal of Pidgin and Creole Languages* 29(1). 49-85.
- Song, Jae Jung. 2010. Word Order Typology. In Song, Jae Jung (a cura di), *The Oxford Handbook of Linguistic Typology*, 253-279. Oxford: Oxford University Press.
- Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Blackwell.
- Stassen, Leon. 2010. The Problem of Cross-Linguistic Identification. In Song, Jae Jung (a cura di), *The Oxford Handbook of Linguistic Typology*, 90-99. Oxford: Oxford University Press.
- Szmrecsanyi, Benedikt & Kortmann, Bernd. 2009. Vernacular universals and angloversals in a typological perspective. In Filppula, Markku & Klemola, Juhani & Paulasto, Heli (a cura di), *Vernacular Universals and Language Contacts: Evidence from varieties of English and Beyond*, 33-53. London – New York: Routledge.
- Tagliamonte, Sali A. 2005. *So who? Like how? Just what?* Discourse markers in the conversations of young Canadians. *Journal of Pragmatics* 37(11). 1896-1915.
- Tagliamonte, Sali A. 2006. *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press.
- Tagliamonte, Sali A. 2012. *Variationist sociolinguistics. Change, observation, interpretation*. Malden: Wiley-Blackwell.
- Trudgill, Peter. 2011. *Sociolinguistic typology*. Oxford: Oxford University Press,
- Tsunoda, Tasaku. 1985. Remarks on transitivity. *Journal of Linguistics* 21. 385-396.

- van der Auwera, Johan. 2013. Semantic maps for synchronic and diachronic typology. In Giacalone Ramat, Anna, Mauri, Caterina & Molinelli, Piera (a cura di), *Synchrony and Diachrony. A dynamic interface*, 153-176. Amsterdam: Benjamins.
- Velupillai, Viveka. 2012. *An introduction to linguistic typology*. Amsterdam: Benjamins.
- Viberg, Åke. 1984. The Verbs of Perception: a Typological Study. *Linguistics* 21. 123-162.
- Weiner, Judith & Labov, William. 1983. Constraints on the agentless passive. *J. Linguistics* 19. 29-58.
- Weinreich, Uriel & Labov, William & Herzog, Marvin I. 1968. Empirical foundations for a theory of language. In Lehmann, Winfred P. & Malkiel, Yakov (a cura di), *Directions for historical linguistics: A symposium*, 95-188. Austin: University of Texas Press.
- Witzlack-Makarevich, Alena. 2019. Argument Selectors. A new perspective on grammatical relations: An introduction. In Witzlack-Makarevich, Alena & Bickel, Balthasar (a cura di), *Argument Selectors: A new perspective on grammatical relations*, 1-38. Amsterdam: Benjamins.
- Wolfram, Walt. 1993. Identifying and interpreting variables. In Preston, Dennis (a cura di), *American dialect research*, 193-221. Amsterdam: Benjamins.
- Wray, Alison & Grace, George W. 2007. The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua* 117(3). 543-578.
- Zeman, Daniel, Nivre, Joakim, Abrams, Mitchell *et al.* 2020. Universal Dependencies 2.6, LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University. Available at: <http://hdl.handle.net/11234/1-3226>. See also <http://universaldependencies.org>.

PARTE SECONDA

RELAZIONI SU INVITO

MASSIMO CERRUTI

Variazione sociolinguistica e processi di grammaticalizzazione

In questo articolo si prende in considerazione una serie di processi di grammaticalizzazione in atto in italiano e nei dialetti italo-romanzi. Concentrandosi sulla variazione interna all'italiano e ai dialetti, si mostra come questi processi procedano a un passo più veloce in alcune varietà di lingua e meno in altre, e si indaga come il contatto intralinguistico possa influire sul passo di grammaticalizzazione delle varietà. Si esamina inoltre l'interazione tra le forze interne a un certo diasistema, capaci di spingere un processo in direzioni diverse, e le forze universali della grammaticalizzazione, in linea di principio unidirezionali; discutendone la rilevanza per lo studio dei processi di grammaticalizzazione in prospettiva sociolinguistica.

Parole chiave: varietà di lingua, passo di grammaticalizzazione, contatto intralinguistico, italiano, dialetti italo-romanzi.

1. Introduzione

Lo studio della variazione sociolinguistica e lo studio dei processi di grammaticalizzazione poggiano su un ampio terreno comune, favorito dalla condivisione di principi epistemologici e dal legame esistente tra fenomeni di variazione e fenomeni di mutamento. Un terreno che tuttavia vede soltanto di rado l'effettiva interazione fra i due ambiti.

In questo contributo si affronteranno gli aspetti di variazione sociolinguistica di alcuni processi di grammaticalizzazione in atto in italiano e nei dialetti italo-romanzi. Concentrandosi sulla variazione interna all'italiano e ai dialetti, e assumendo quale categoria d'analisi fondamentale la nozione di varietà di lingua, si mostrerà come certi processi si collochino a uno stadio più avanzato, e quindi procedano con un passo più veloce, in alcune varietà del sistema e meno in altre.

Più in particolare, fatta qualche considerazione preliminare (§ 2), si discuterà come le varietà spontanee (sub-standard) siano tipi-

camente più grammaticalizzate della varietà codificata (standard) (§ 3) e come le dinamiche di convergenza ‘verticale’ interne a un sistema possano influire sul passo di grammaticalizzazione delle sue varietà (§ 4); confrontando negli stessi termini lingua standard e dialetto (§ 5). Si trarranno infine alcune impressioni di carattere generale sulla rilevanza dello studio dei processi di grammaticalizzazione in prospettiva sociolinguistica (§ 6).

2. Considerazioni preliminari

Come già accennato, lo studio della variazione sociolinguistica e lo studio dei processi di grammaticalizzazione condividono una serie di principi retrostanti, comuni del resto a più ambiti di studio. Si possono citare, fra gli altri: (i) l’attenzione prioritaria rivolta all’uso della lingua, così come in tutti gli approcci *usage-based* praticati nelle scienze del linguaggio (v. ad es. Bybee 2011; Torres Cacoullous 2011); (ii) la convinzione che sia il parlato l’habitat naturale dei fatti da osservare (v. ad es. Kortmann & Schneider 2011; Couper-Kuhlen 2011; Hymes 2020); (iii) una concezione delle categorie basata sulla gradualità e non sulla categoricità, per cui i confini tra i fenomeni sono in linea di principio sfumati (v. ad es. Traugott & Trousdale 2010; Giacalone Ramat & Mauri 2012; Seiler 2018); (iv) l’importanza che ha, per la definizione stessa dei fenomeni, circoscriverne il contesto (v. ad es. Heine 2002; Diewald 2002; Poplack 2011); e (v) il principio per cui l’oggetto fondamentale di studio sono i comportamenti linguistici di una collettività, anziché i fatti di comportamento e produzione individuale (“there is no individual usage act or event which constitutes the grammaticalization of a linguistic form”, Harder & Boye 2011: 59; v. anche Milroy & Milroy 1997; Labov 2007).

Questi (e altri) principi si innestano evidentemente sullo stretto rapporto che esiste tra fenomeni di variazione e fenomeni di mutamento, per cui i primi in parte prefigurano i secondi e in parte ne sono il risultato (v. anche Croft 2010; Giacalone Ramat, Mauri & Molinelli 2013).

Per quanto riguarda poi, più specificamente, la grammaticalizzazione, la variazione interna a una lingua è essa stessa un elemento costitutivo del processo; gli stadi di un processo di grammaticalizzazione rappresentano infatti, allo stesso tempo, fasi diacroniche di un muta-

mento e manifestazioni sincroniche della variabilità di una lingua (“a grammaticalization scale accounts for both synchronic variation and diachronic change”, Lehmann 1985: 6). È noto invero, ed è cruciale per quanto si dirà più avanti, che stadi diversi di un processo possono coesistere in sincronia; uno stesso processo può quindi trovarsi in fasi più avanzate in una o più varietà di una lingua e meno in altre (v. ad es. Seiler 2018).

La variazione intralinguistica contribuisce inoltre a definire certi parametri di grammaticalizzazione; ad esempio, i parametri che Lehmann (1985: 4-5) chiama *obligatorification* e *fixation* sono determinati da una riduzione progressiva di variabilità, rispettivamente paradigmatica e sintagmatica (ma v. Heine & Narrog 2010). E si può aggiungere che la variazione interna a una lingua è correlata ad almeno alcuni principi di grammaticalizzazione; il principio del *layering* (Hopper 1991: 22-23), ad esempio, identifica tipicamente un sottoinsieme delle manifestazioni della variazione sociolinguistica (“layering [...] is a subset of what variationists refer to as inherent variability”, Poplack 2011: 210).

Ciò nondimeno, e sia pure a fronte di un interesse crescente per gli approcci ‘integrati’ allo studio della variazione grammaticale (v. ad es. Kortmann 2004; Cornips & Corrigan 2005; Léglise & Chamoreau 2013), è ancora tutto sommato raro che i processi di grammaticalizzazione siano esplorati in prospettiva sociolinguistica (v. ad es. Nevalainen & Palander-Collin 2011: 118; Poplack *et al.* 2018); vuoi per una certa differenza di angolatura, od ottica, più generale fra ‘linguistica interna’ e ‘linguistica esterna’, vuoi per motivi di carattere più specifico. Uno di questi è che, per sua natura, la grammaticalizzazione dà luogo a elementi linguistici con una salienza discorsiva secondaria, se non privi di salienza discorsiva (Harder & Boye 2011: 63)¹, e quindi meno inclini di altri a sviluppare significati sociali (v. ad es. Meyerhoff 2013: 32-35). Un altro è che la grammaticalizzazione, come tipo di mutamento in cui “form and meaning pairings change” (Traugott 2011: 21), investe elementi linguistici interessati da un cambiamento di forma e significato; e la sociolinguistica è tradizionalmente riluttante ad affrontare elementi come questi, portatori di significato autonomo, poiché ‘eccentrici’ rispetto

¹ Occorrerebbe però discutere, a questo proposito, della collocazione dei fenomeni di pragmaticalizzazione (per cui si vedano ad esempio Diewald 2011 e Giacalone Ramat 2017: 6-8).

alla nozione di variabile sociolinguistica (che, notoriamente, è fondata sull'equivalenza semantica e funzionale di varianti formali; la questione è discussa ad esempio in Cerruti 2011a).

Non a caso, la maggior parte degli studi variazionisti sulla grammaticalizzazione si concentrano su fatti associati al *layering* (v. ad es. Tagliamonte 2000; Poplack 2011; Torres Cacoullous 2011; Arroyo & Vellón Lahoz 2018), connessi quindi essenzialmente alla coesistenza di forme diverse con una stessa funzione; e perciò più immediatamente inquadrabili nel paradigma delle variabili sociolinguistiche. Soltanto sporadici (v. ad es. Pichler 2016) sono invece gli studi su fatti di variazione sociolinguistica determinati da principi quali *divergence* e *de-categorialization* (nel senso di Hopper 1991), che, inevitabilmente, richiedono di confrontarsi più da vicino con le diverse funzioni assunte da una stessa forma.

3. *Il passo della grammaticalizzazione*

Alcuni studi variazionisti si inscrivono nell'ambito di ricerca del cosiddetto 'passo' della grammaticalizzazione; sono cioè ispirati all'ipotesi che una serie di processi di grammaticalizzazione si trovi a uno stadio più avanzato, e quindi proceda con un passo più veloce, in una lingua (o in alcune lingue) di una certa famiglia e meno in altre (cfr. Lamiroy 2011; Lamiroy & De Mulder 2011: 302; Carlier, De Mulder & Lamiroy 2012). In quest'ambito, un lavoro recente condotto in prospettiva variazionista è Poplack *et al.* (2018), che affronta l'evoluzione del congiuntivo da modo verbale indicante la non-fattualità a semplice marca di subordinazione, v. (1).

- (1) *making a semantic contribution > being an obligatory concomitant of subordination*
(Bybee, Perkins & Pagliuca 1994: 214)

Nello studio citato si confrontano quattro corpora di parlato spontaneo, in larga parte colloquiale, di lingue romanze; si pone in relazione, per ciascun corpus, la frequenza d'uso del congiuntivo con alcuni parametri di grammaticalizzazione; e, sulla base di questi valori, si commisura lo stadio di evoluzione del congiuntivo tra i diversi corpora².

² L'equivalenza delle costruzioni originarie è tuttavia un'ipotesi euristica, che poggia su "an *idealized* state of the ancestor in which the subjunctive 1) made a semantic

I risultati dello studio sono ampiamente coerenti con l'ipotesi della scala di grammaticalizzazione delle lingue romanze, o *Romance grammaticalization cline*, v. (2); in base alla quale gli stadi meno avanzati di una serie di processi di grammaticalizzazione si avrebbero in spagnolo e gli stadi più avanzati in francese, con l'italiano in posizione intermedia:

- (2) *French > Italian > Spanish*³
(Carlier, De Mulder & Lamiroy 2012: 288)

È importante sottolineare, tuttavia, che in Poplack *et al.* (2018) l'analisi si concentra sul parlato colloquiale; pone quindi a confronto non tanto lingue diverse quanto varietà di una stessa natura in lingue diverse. Secondo invece l'ipotesi di *cline* avanzata per le lingue romanze, così come per scale analoghe relative a lingue di altri rami⁴, a essere più e meno grammaticalizzate sarebbero le lingue nel loro complesso. Il che può essere problematico, specie considerando che stadi diversi di un processo possono coesistere in sincronia e uno stesso processo può dunque trovarsi in fasi più avanzate in certe varietà di lingua e meno in altre (cfr. § 2). Da più parti, infatti, si è osservato che all'interno di una stessa lingua classi di varietà di natura differente spesso non procedono allo stesso passo; è noto ad esempio che molti processi di grammaticalizzazione prendono avvio e avanzano nel parlato colloquiale, mentre tendono a essere rallentati dalla codificazione nella varietà che ne è soggetta (v. ad es. Laitinen 2004; Miller 2004; Berruto 2007; Kortmann & Schneider 2011; Kühl & Braunmüller 2014; Haspelmath & Michaelis 2017)⁵.

contribution, and 2) did so both consistently and categorically, i.e. whenever the conditions for its use were present" (Poplack *et al.* 2018: 223).

³ Qui il segno ">" è da intendere a tutti gli effetti come simbolo di maggiore, nel senso di "x è più grammaticalizzato di y".

⁴ Per le lingue germaniche, ad esempio, si è proposta una scala di questo tipo: inglese > neerlandese > lingue scandinave > tedesco (cfr. Vismans *et al.* 2010; Lamiroy 2011: 170).

⁵ A questo proposito, può essere opportuno esplicitare che l'ipotesi del passo della grammaticalizzazione non assume né che le lingue o varietà di lingua in fase iniziale siano scarsamente grammaticalizzate, né che queste col tempo pervengano al grado di grammaticalizzazione più avanzato. I processi di grammaticalizzazione non giungono inesorabilmente a compimento: possono arrestarsi a un certo stadio in una lingua o varietà di lingua e conoscere uno sviluppo ulteriore in un'altra.

Parrebbe a questo punto meno problematico delineare una scala di grammaticalizzazione di varietà di lingua; secondo la quale, come in (3), le varietà spontanee (sub-standard) sarebbero tipicamente più grammaticalizzate, procederebbero cioè a un passo più veloce, della varietà codificata (standard)⁶:

- (3) varietà spontanee (sub-standard) > varietà codificata (standard)

Ad esempio, per quanto riguarda ancora le lingue romanze, proprio l'uso del congiuntivo risulta più grammaticalizzato nelle varietà spontanee, sub-standard, che nella varietà codificata⁷ (cfr. Poplack *et al.* 2018: 219-221). In merito all'italiano, in particolare, Digesto (2019) mostra come la scelta del congiuntivo nel parlato colloquiale dipenda prevalentemente da fattori di carattere sintattico e lessicale, ossia dalla struttura della subordinata e dal verbo reggente; anche quando, come in (4) e (5), lo stato di cose descritto dalla subordinata è presentato come reale. Di contro, per buona parte delle grammatiche normative l'uso del congiuntivo è ancora primariamente correlato alla semantica non-fattuale della subordinata (cfr. Prandi 2018: 26-31; Renzi 2019: 19-28; Digesto 2019: 16-39). Beninteso, è da ascrivere allo stesso processo la sostituzione dell'indicativo al congiuntivo, come in (6), giacché “after the generalization of subordinating uses, the only further development for such grams is their gradual loss from the language” (Bybee, Perkins & Pagliuca 1994: 213-214).

- (4) *io lo leggo sono su Facebook davvero ma il fatto che io le legga su Facebook la dice lunga*
(KIP, www.kiparla.it, conversazioni libere)
- (5) *non dovrebbe essere Presidente del Consiglio però è bello che lo sia*
(KIP, www.kiparla.it, conversazioni libere)
- (6) *ma non credo che sono state cambiate un po' di cose*
(C-ORAL-ROM, <http://lablita.dit.unifi.it/coralrom>; Digesto 2019: 73)

⁶ Fatti salvi, ovviamente, alcuni casi eccezionali. Ne è un esempio il giapponese, che presenta una serie di processi di grammaticalizzazione più avanzati nella varietà codificata (cfr. Narrog & Ohori 2011).

⁷ Oltre che, come si diceva sopra, più grammaticalizzato nelle varietà spontanee di certe lingue e meno in quelle di altre.

Si può portare come ulteriore esempio il processo per cui i clitici oggetto tendono a evolvere in direzione di affissi di accordo verbale, tracciato in (7). Anche in tal caso, nelle lingue romanze sono le varietà spontanee, sub-standard, a ospitare gli stadi più avanzati del processo: i clitici oggetto, che nella varietà codificata sono usati per riferimenti deittici o anaforici deenfatici, in varietà spontanee, sub-standard (Bossong 2003: 38), possono comparire anche all'interno di frasi nelle quali sia già presente un nominale coreferenziale, comportandosi così in modo analogo a marche di accordo fra il verbo e l'oggetto (v. anche Lehmann 1982: 233-241; Corbett 2006: 99-113).

- (7) *free pronoun > clitic pronoun > variable affix > invariable affix*
(Lehmann 1982: 239)

Per ciò che concerne l'italiano, è noto ad esempio come le dislocazioni a destra interrogative tendano a perdere marcatezza nel parlato colloquiale (Berretta 2002: 183). In particolare, con referenti discorsivamente nuovi ma accessibili su base contestuale (cfr. Crocco & Badan 2020), come in (8), il clitico oggetto viene a perdere valore anaforico ed è portato a co-occorrere con un nominale coreferenziale integrato intonativamente nella stessa frase (o quanto meno interpretabile come tale). Inoltre, specialmente in varietà popolari, la presenza di un clitico oggetto tende a essere categorica in certi contesti (cfr. Berretta 1989: 142-148). È il caso ad esempio di dislocazioni a sinistra con un pronome tonico preverbale, come in (9), in cui il nominale coreferenziale è oltretutto integrato sintatticamente, per la presenza di accusativo preposizionale. Si può aggiungere che le varietà popolari sono le più sensibili all'influenza dei dialetti di sostrato; e in questi, non a caso, lo stesso processo di grammaticalizzazione parrebbe trovarsi in fasi molto avanzate (cfr. Cerruti 2008).

- (8) *lo vuole un caffè?*
(Berruto 2012: 77)
- (9) *a me, come tutti i braccianti agricoli delle Basse, mi hanno
destinato in fanteria*
(Berretta 1989: 146)

Tuttavia, casi di questo tipo rendono opportuno un *caveat*: varietà codificata e varietà standard non sempre si equivalgono. La varietà codificata, tradizionale, descritta nelle grammatiche normative, è infatti spesso improntata alla lingua delle opere letterarie, o comunque all'uso scritto

e formale, e può risentire delle ideologizzazioni proprie di una certa tradizione di pensiero, talvolta lontane dalla realtà dei testi. Accanto a questa varietà, però, consolidatasi in Età moderna, si assiste oggi in molte lingue allo sviluppo spontaneo di una varietà standard d'uso medio, neo-standard: non più modellata sull'uso scritto e formale, scarsamente codificata e contrassegnata dalla presenza di tratti tipici di varietà spontanee, sub-standard (v. ad es. Auer 2017a). L'affermarsi di una varietà standard come questa, particolarmente ricettiva nei confronti di tratti sub-standard, comporta dunque un progressivo avvicinamento tra gli usi di norma e il parlato colloquiale. In italiano, ad esempio, considerando processi come (1) e (7), il passo di grammaticalizzazione della varietà neo-standard parrebbe per certi versi effettivamente avvicinarsi a quello di varietà sub-standard (cfr. § 4.2).

Al contempo, nemmeno lo spazio di variazione del sub-standard è uniforme: il parlato colloquiale di colti, a cui si riferiscono tra gli altri gli enunciati (4) e (5), ha difatti natura e caratteri linguistici almeno in parte diversi dalle varietà di semicolti. L'italiano popolare, ad esempio, per molti parlanti dialettofoni non è una varietà spontanea nativa; risente, anche per questo, più di altre varietà sub-standard del contatto con i dialetti di sostrato⁸; e ha le caratteristiche di una varietà obsolescente (cfr. Berruto 2014). Non a caso, come si discuterà più avanti, può accadere che uno stesso processo abbia nel parlato colloquiale uno stadio di avanzamento, e quindi un passo, differente da quello che presenta in italiano popolare (cfr. §§ 4.1 e 4.2).

4. *Il contatto fra varietà di lingua*

Il contatto fra lingue, a cui si è appena accennato in relazione all'influenza dei dialetti sull'italiano, può notoriamente innescare, oltre che accelerare, un processo di grammaticalizzazione (cfr. Heine & Kuteva 2005); ed è stato ampiamente esplorato in questa prospettiva (v. anche Matras 2011). Meno studiato è il ruolo esercitato dal

⁸ In casi specifici, proprio il contatto coi dialetti può dar luogo a esiti eccezionali rispetto alla scala in (3). Ne è un esempio la realizzazione di congiunzioni subordinanti con costituenti separati, sul modello del dialetto di sostrato; come in *per non che rida* o *fin non che è nata* (Cerruti 2011b: 178) in varietà sub-standard di area settentrionale (Berruto 2009: 33). In questo caso, infatti, sono evidentemente più grammaticalizzati i corrispondenti univerbati *perché* e *finché* della varietà standard.

contatto fra varietà di una stessa lingua, che pure può contribuire al passo, se non all'innescò, di un processo di grammaticalizzazione. Un certo processo, una volta che è innescato (per contatto o per altri fattori) e prende l'avvio in una varietà di una lingua (si pensi ad esempio all'italiano popolare in molti casi di grammaticalizzazione indotti dal contatto col dialetto; cfr. §§ 3 e 4.1), procede infatti con un passo che dipende anche dalla rete di relazioni, e dunque dal contatto, fra le varietà di quella lingua. E se il contatto fra varietà ha effetti su una serie di processi, può influire sul passo di grammaticalizzazione complessivo delle varietà interessate.

A questo proposito, la scala di grammaticalizzazione in (3) porta a ritenere che la convergenza 'verso il basso' – cioè della varietà codificata (standard) in direzione delle varietà spontanee (sub-standard), v. (10) – acceleri il passo della varietà meno grammaticalizzata, ovvero della varietà codificata, standard; e che la convergenza 'verso l'alto' – cioè delle varietà spontanee (sub-standard) in direzione della varietà codificata (standard), v. (11) – rallenti il passo delle varietà più grammaticalizzate, ossia delle varietà spontanee, sub-standard:

(10) varietà spontanee (sub-standard) ← varietà codificata (standard)

(11) varietà spontanee (sub-standard) → varietà codificata (standard)

Si tratta evidentemente di due dinamiche di verso contrario, che sono alla base di tendenze di carattere più generale: la convergenza 'verso il basso' è ad esempio il principale motore dello sviluppo spontaneo di varietà neo-standard (cfr. § 3), e la convergenza 'verso l'alto' è una delle cause dell'obsolescenza dei tratti socio-geograficamente più marcati e delle varietà popolari nel loro complesso (v. ancora § 3); due tendenze osservabili non soltanto in italiano ma in varie altre lingue europee (cfr. Auer, Hinskens & Kerswill 2005; Auer 2017b). Il passo della grammaticalizzazione di alcuni processi, e di certe varietà di lingua, può quindi essere interpretato anche come esito di queste tendenze.

4.1 Mica

Consideriamo ad esempio il processo di grammaticalizzazione che interessa la particella *mica*, da negazione non canonica a negazione canonica. Il percorso delineato in (12) è ricavato da Squartini (2017: 212), che riconosce tre primi stadi di sviluppo della particella, ovvero tre manifestazioni via via più grammaticalizzate di negazione non canonica, e un

ultimo stadio in cui la particella acquisisce il valore di negazione canonica. La semantica di questi quattro stadi è definita essenzialmente dal tipo di contenuto negato, se dato o nuovo rispetto al contesto discorsivo (*discourse-old*, *discourse-new*) e alle conoscenze condivise dall'ascoltatore (*hearer-old*, *hearer-new*).

- (12) (i) > (ii) > (iii) > (iv)
discourse-old discourse-old discourse-new discourse-new
hearer-old hearer-new hearer-old hearer-new
 (Squartini 2017: 212)

I primi tre stadi, in cui la negazione ha un valore pragmatico, sono propri dell'uso standard di *mica* (come si può desumere da trattazioni grammaticali dell'italiano e lavori specifici sul tema; v. ad es. Cinque 1976; Serianni 1989: 428; Manzotti & Rigamonti 1991; Bernini & Ramat 1996: 17ss.). È ad esempio rappresentativo del terzo stadio il caso (13), in cui *mica* nega un contenuto nuovo rispetto al contesto discorsivo ma presentato come dato in relazione alle conoscenze condivise dall'ascoltatore (e dal lettore). I soli usi di *mica* come negazione canonica parrebbero riscontrarsi nelle varietà popolari di alcuni italiani regionali settentrionali (cfr. Ramat 2006; Mosegaard Hansen & Visconti 2012), essenzialmente per effetto del contatto con i dialetti di sostrato. In vari dialetti settentrionali, infatti, la particella corrispondente a *mica* ha la stessa semantica dell'italiano *non*: può comparire anche in contesti di negazione non canonica ma ha raggiunto a tutti gli effetti il quarto stadio del processo; come [*'mia*] nell'enunciato bresciano in (14).

- (13) *Etruschi: la tesi di uno studioso controverso. [...]. S'aggiusta il colletto, Giovanni Semerano, ti fa segno di avvicinarti, a ottantott'anni non è abituato a strillare, e con fare sornione ti dice: "Il vero mistero dell'etrusco non è mica la lingua, sa. Sono gli studiosi"*
 (CORIS, http://corpora.dslo.unibo.it/coris_ita.html, periodici)

- (14) [ma al tɔ te'dɛhk 'ando 'ɛ-l?
 ma il tuo tedesco dove è-3SG.SBJ
 al dih al vu'lia 'mia jni...
 3SG.SBJ dice 3SG.SBJ voleva NEG venire
 a'lyra l o ko'pat]
 allora l'ho accoppato
 "ma il tuo tedesco dov'è? ha detto che non voleva venire...
 allora l'ho accoppato"
 (ParVa, www.mediling.eu; Cerruti 2018: 94)

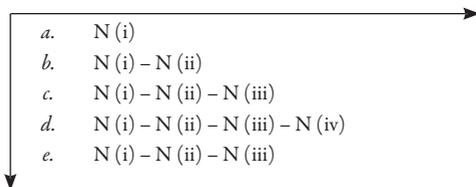
È necessario osservare, però, che gli usi di *mica* come negazione canonica parrebbero fortemente in regresso. Ad esempio in un corpus come il ParVa (da cui è tratto fra l'altro il brano 14), che è costituito da interviste a parlanti anziani dialettofoni e ha ampie porzioni in italiano popolare bresciano (cfr. Guerini 2016), non vi sono occorrenze di *mica* in contesti di negazione canonica (cfr. Cerruti 2018: 94; v. anche Ballarè 2020a: 139-140, che registra lo stesso dato in un altro corpus di parlato di area lombarda). Il venir meno degli usi di *mica* come negazione canonica, già circoscritti a varietà popolari settentrionali, rappresenterebbe perciò la perdita di un tratto socio-geograficamente marcato (indotta con ogni probabilità sia da fattori linguistici sia da fattori extralinguistici; per una prima discussione v. ad es. Cerruti 2020: 132-143).

In questa prospettiva, dunque, lo sviluppo di *mica* da negazione non canonica a negazione canonica sarebbe innescato dal contatto interlinguistico, in particolare dalla convergenza di varietà popolari settentrionali di italiano verso i dialetti di sostrato; e potrebbe rappresentare un caso di grammaticalizzazione indotta dal contatto (si veda però più avanti). Il successivo retrocedere del processo, con il venir meno della funzione di negazione canonica di *mica*, sarebbe invece influenzato dal contatto intralinguistico, ovvero dalla convergenza delle stesse varietà popolari verso lo standard. Segnatamente, quest'ultimo passaggio potrebbe essere inteso come un caso di 'ritrazione' (*retraction* in Haspelmath 2004), indotto appunto dal contatto fra varietà di lingua⁹.

Nella Figura 1, che riprende lo schema con cui Haspelmath (2004) illustra i fenomeni di 'ritrazione', i numeri romani possono esser fatti corrispondere ai quattro stadi del processo di grammaticalizzazione tracciato in (12); in questi termini, il passaggio da *c* a *d* verrebbe a rappresentare la fase di grammaticalizzazione indotta dal contatto, e il passaggio da *d* a *e* la fase di 'ritrazione', provocata appunto dalla convergenza delle varietà popolari verso lo standard.

⁹ Fra gli altri, si possono citare come esempi di ritrazione la perdita delle costruzioni impersonali con derivati di *homo* in varietà italo-romanze antiche (Giacalone Ramat & Sansò 2007) e il retrocedere recente della perifrasi *aviri a* + infinito in siciliano da forma di futuro a marca di modalità (Amenta 2020: 95-108).

Figura 1 - 'Ritrazione' (Haspelmath 2004: 33, con adattamenti)



Il caso di *mica* mostra altresì come siano in gioco forze differenti, che seguono percorsi naturalmente diversi nelle loro motivazioni e che possono perciò spingere uno stesso processo in direzioni dissimili: da un lato le forze 'interne', universali, della grammaticalizzazione, che sono in linea di principio unidirezionali, e dall'altro le forze 'esterne', locali, proprie di un certo diasistema, che possono indurre un processo tanto ad avanzare quanto a retrocedere. Un mutamento, del resto, non ha sempre un andamento lineare e proprio la variazione sociolinguistica può essere responsabile di alcune fluttuazioni (v. ad es. Labov 2001 e Janda 2001; cfr. Seiler 2018).

Nel nostro caso tuttavia, in assenza di riscontri empirici in diacronia sulle varietà in questione, per l'innescò del processo non è possibile escludere un'interpretazione diversa dalla grammaticalizzazione per contatto; un'interpretazione per la quale si può chiamare in causa la nozione di "polysemy copying [...], or calquing or loan translation" (Heine & Kuteva 2005: 100). In questa prospettiva, cioè, il valore di negazione canonica che ha *mica* in alcune varietà popolari settentrionali sarebbe comunque da imputarsi al contatto con i dialetti di sostrato ma non rappresenterebbe l'esito di un processo di grammaticalizzazione; sarebbe invece dovuto al trasferimento diretto in italiano della semantica che ha la particella corrispondente in dialetto. Per riprendere l'esempio bresciano, la semantica di ['mia] sarebbe quindi trasferita direttamente al corrispondente *mica* dell'italiano popolare, senza il verificarsi di un vero e proprio passaggio di *mica* da negazione non canonica a negazione canonica. Ciò che distingue, infatti, la grammaticalizzazione per contatto dalla 'copia di polisemia', o più in generale da fenomeni di calco, è appunto la presenza vs. assenza di stadi intermedi di sviluppo (cfr. Heine & Kuteva 2005: 102).

Da un'angolatura più spiccatamente variazionista, o meglio 'variazionale' (poiché incentrata sulla concezione di una lingua come somma di varietà), si può poi aggiungere che per l'effettiva esistenza di un processo occorre riscontrare stadi contigui di sviluppo quanto

meno in una delle varietà sociali, o socio-geografiche, di una lingua. Per poter affermare, ad esempio, che la particella *mica* muti la propria funzione da negazione non canonica a negazione canonica, e che quindi il secondo valore si sviluppi a tutti gli effetti a partire dal primo, è necessario appurare che questo passaggio avvenga per lo meno in una stessa varietà sociale, o socio-geografica, d'italiano; altrimenti, più che di fasi di un processo potrebbe trattarsi di esiti indipendenti, diffusi l'uno presso un certo gruppo sociale e l'altro presso un gruppo sociale differente (poniamo, l'uso di *mica* come negazione non canonica in parlanti colti e di *mica* come negazione canonica in parlanti semicolti). Il che, del resto, è coerente con il principio per cui “grammaticalization [...] can only be captured at the collective level” (Harder & Boye 2011: 59; su questi aspetti v. anche Laitinen 2004; Miller 2004; Nevalainen & Palander-Collin 2011; Seiler 2018: 83; cfr. § 2).

4.2 Pronomi relativi

In questi stessi termini si possono discutere alcuni fenomeni di variazione osservabili nel settore delle costruzioni relative. Consideriamo ad esempio il processo di grammaticalizzazione che porta un pronome relativo a evolvere in un introduttore invariabile di frase relativa (e da questo, eventualmente, in un complementatore; come in 15).

- (15) *relative > complementizer*
(Heine & Kuteva 2002: 254)

Nell'italiano di parlanti colti, specialmente in produzioni formali, compare il pronome relativo del tipo *il quale*, proprio della varietà standard, che si accorda in genere e numero con la testa della relativa ed esprime la funzione sintattica dell'elemento relativizzato, v. (16) e (17). La forma *il quale* è presente anche in italiano popolare, ma in questa varietà si comporta sostanzialmente da introduttore invariabile di frase relativa: non si accorda con l'elemento relativizzato e non ne esprime la funzione sintattica; se non, alla stregua di *che*, in combinazione con un clitico, v. (18) e (19); cfr. Fiorentino (1999: 117-118), Cerruti (2017: 68-69).

- (16) *Pirandello insiste moltissimo anche nelle sue lettere a Gandusio, il quale non capisce niente di tutto ciò*
(KIP, www.kiparla.it, lezioni)

- (17) *l'ultima parte del corso [...] quella sul sessantotto, alla quale eh dedicherò le ultime tre lezioni*
(KIP, www.kiparla.it, lezioni)
- (18) *ci mettono le noci, il quale si seccheranno bene*
(Alisova 1965: 329)
- (19) *la maestra mandò a chiamare Peppone, il quale gli ["al quale"] disse: quando sono morta...*
(Alisova 1965: 329)

Perché possa effettivamente trattarsi di un processo di grammaticalizzazione, questi due diversi comportamenti di *il quale* dovrebbero corrispondere a due stadi di sviluppo differenti osservabili per lo meno in una stessa varietà sociale; oppure dovrebbe essere possibile individuare stadi intermedi, ascrivibili comunque (se non altro, insieme a uno degli stadi precedenti) ad almeno una delle varietà in questione¹⁰. Altrimenti, l'uso del pronome relativo originario come introduttore invariabile di frase relativa potrebbe rappresentare non una fase di un processo di grammaticalizzazione ma un esito indipendente, di altra natura, venutosi a creare con il mero trasferimento della forma *il quale* in italiano popolare e circoscritto a un gruppo di parlanti che di quella forma, sostanzialmente, non fa altro uso.

Abbiamo invece argomenti per ricondurre effettivamente a stadi di un processo le diverse manifestazioni sincroniche della variabilità di *dove*; la cui sovraestensione d'impiego, a partire dall'uso come relativizzatore di locativi, parrebbe seguire un percorso di grammaticalizzazione descrivibile in questi termini:

- (20) *locative > concern > subject*
(Ballarè & Inglese 2021)

In una prima fase, *dove* ha appunto la funzione di relativizzare un elemento con valore locativo, come in (21); questo è il comportamento conforme all'uso standard, non marcato, riscontrabile in linea di

¹⁰ Qualche indizio a riguardo si potrebbe ricavare da esempi come *quei paesi nel quale vi era una forte immigrazione* e *i problemi del quale mi sto lamentando io*, tratti il primo da una tesi di laurea triennale (Ballarè, c.p.) e il secondo da un'intervista a un parlante giovane con istruzione professionale (ParlaTO, www.kiparla.it), in cui *il quale* non si accorda con l'elemento relativizzato ma ne esprime la funzione sintattica attraverso una preposizione. Allo stato attuale, tuttavia, l'entità dei materiali a nostra disposizione non è sufficiente ad avanzare ipotesi.

principio in tutte le varietà dell'italiano. In una fase successiva, *dove* sviluppa un carattere sostanzialmente logico-testuale di *concern* (o *aboutness*), come in (22), “in which *dove* cannot be characterized as relativizing a specific syntactic role, but rather acts as a more general linker between the main clause and the subordinate clause” (Ballarè & Inglese 2021: 18-19; cfr. Heine & Kuteva 2002: 201). Questa seconda fase è propria di varietà spontanee, sub-standard, di italiano, fra cui italiano colloquiale e italiano popolare; ma caratterizza altresì alcuni usi neo-standard: con valore di *concern*, infatti, *dove* compare anche nella prosa giornalistica e in altri ‘testi modello’ (nel senso di Ammon 2003; cfr. Ballarè & Micheli 2018). A partire, poi, da contesti come (22), “in which what is relevant is the topic-comment relationship between the main clause and the subordinate” (Ballarè & Inglese 2021: 25), *dove* sviluppa la possibilità di relativizzare il soggetto di frase (è noto, del resto, come marche di *topic* tendano a evolvere in marche di soggetto). In una terza fase, quindi, *dove* è usato per la relativizzazione di un soggetto; come nell'esempio (23)¹¹. Lo stadio più avanzato del processo parrebbe tuttavia riscontrarsi soltanto in varietà ‘basse’, e in modo particolare nel parlato di semicolti (cfr. Ballarè *et al.* 2019: 85-86); a conferma di come le varietà sub-standard (o per lo meno alcune di queste, cfr. § 3) siano più grammaticalizzate, procedano cioè a un passo più veloce, della varietà standard.

- (21) *è un posto dove posso dormire, nessuno mi disturba però non è camera mia*
(KIP, www.kiparla.it, interviste semistrutturate)
- (22) *chiaramente la FIAT è un po' come quelle madri un po' oppressive dove non ti manca mai nulla*
(ParlaTO, www.kiparla.it, interviste semistrutturate; Ballarè & Inglese 2021: 19)
- (23) *lati negativi di Venaria [...] i ragazzini under sedici dove stan lì a far casino al giardinetto sotto casa*
(ParlaTO, www.kiparla.it, interviste semistrutturate; cfr. Ballarè & Inglese 2021: 22)

¹¹ E da qui l'uso di *dove* come relativizzatore può poi sovraestendersi a posizioni via via più a destra della gerarchia di accessibilità del sintagma nominale (come accade ad esempio in inglese, cfr. Brook 2011; v. anche Murelli 2011: 184).

Al contempo, si può ritenere che la diffusione nel neo-standard dell'uso di *dove* come marca di *concern* (v. sopra) sia un effetto della convergenza degli usi di norma 'verso il basso', ossia verso le varietà spontanee, sub-standard. La stessa dinamica può essere chiamata in causa per almeno altri due processi a cui si è fatto cenno: lo sviluppo del congiuntivo da modo verbale indicante la non-fattualità a marca di subordinazione e l'evoluzione dei clitici oggetto in direzione di affissi di accordo verbale; è in varietà sub-standard che questi due processi si collocano agli stadi più avanzati (cfr. § 3), ed è proprio verso questi stadi che sembrerebbe orientarsi lo sviluppo degli stessi processi in italiano neo-standard (v. ad es. Berruto 2012: 80-81; Renzi 2019: 19-28; Ballarè 2020b: 478-482).

5. *Lingue standard e dialetti*

Tornando all'ipotesi della scala di grammaticalizzazione delle lingue romanze (cfr. § 3), è opportuno aggiungere che gli studi condotti in quest'ambito hanno preso in considerazione soprattutto lingue standard, riservando scarsa attenzione ad altri tipi funzionali di lingua, quali dialetti e lingue minoritarie (fanno eccezione lavori come Lamiroy & Pineda 2017 e Miola 2017, 2018; v. anche Haspelmath & Michaelis 2017). Lingue come queste, usate principalmente nel parlato e per definizione scarsamente codificate (si veda però Regis 2014), rivestono tuttavia interesse per l'ipotesi citata, giacché, essendo soggette soltanto debolmente all'azione uniformante di regole normative, potrebbero mostrarsi nel complesso più grammaticalizzate delle lingue standard.

Alcune regolarità parrebbero in effetti riscontrabili, se non altro in relazione a uno stesso sottogruppo di lingue. Per quanto riguarda ad esempio l'italoromania, il confronto tra dialetti e italiano rivela come un certo numero di processi di grammaticalizzazione si collochino a uno stadio più avanzato, e quindi procedano con un passo più veloce, nelle varietà dialettali; o per lo meno in alcune di queste (cfr. Miola 2017, 2018; v. anche § 3). Ciò nondimeno, come sottolineato in Miola (2018: 131-132), la natura di dialetto (o lingua minoritaria) in sé "non sembra implicare necessariamente l'occupazione di una posizione più grammaticalizzata di quella delle lingue standard". Vari fattori sociali, oltre che linguistici, possono infatti influire sul passo

di grammaticalizzazione di una lingua, accelerandolo oppure rallentandolo; fra gli altri, il tipo e la gamma di situazioni di contatto in cui è coinvolta la lingua, il suo grado di standardizzazione e scritturalizzazione, le dimensioni e il tipo di reti sociali della sua comunità parlante (cfr. Milroy & Milroy 1985; De Mulder & Lamiroy 2012; Miola 2018).

Occorre poi tener presente, come già osservato, che le varietà di una stessa lingua non procedono sempre al medesimo passo; uno stesso processo può appunto collocarsi in fasi più avanzate in certe varietà e meno in altre. Nei paragrafi precedenti la questione è stata affrontata in relazione all'italiano ma bisogna ricordare che anche un dialetto è tipicamente articolato al proprio interno in varietà; per ciò che concerne il dominio italoromanzo, in alcune aree il *continuum* dialettale si caratterizza ad esempio per la compresenza di numerose varietà locali ('basse') e una varietà di koinè ('alta', di prestigio), verso la quale le prime tendono a convergere (v. ad es. Regis 2018).

Ci si potrebbe chiedere altresì se, e in quali termini, si pongano per i dialetti le stesse questioni che sono state discusse per le lingue standard; ovvero, se alcune varietà di dialetto siano generalmente più grammaticalizzate di altre e se la convergenza verticale, che in questi *continua* tende ad essere unidirezionale ('verso l'alto'), accelera oppure rallenti il passo delle varietà locali. Dei dialetti italoromanzi è stata in realtà poco studiata la variazione interna e, *a fortiori*, l'esistenza di differenze nel passo di grammaticalizzazione delle varietà. L'impressione che comunque si ricava, allo stato attuale delle conoscenze, è che non siano sempre le stesse varietà di dialetto ad essere le più o le meno grammaticalizzate: essenzialmente, di certi processi gli stadi più avanzati si avrebbero nelle varietà locali e di altri nella varietà di koinè.

Per quanto riguarda ad esempio il piemontese, si può notare che nella varietà torinese, ossia di koinè, la selezione dell'ausiliare segue fondamentalmente un *pattern* di intransitività scissa (come in italiano), mentre in alcune varietà locali, specialmente del biellese e del Piemonte meridionale, *avere* ha una distribuzione più ampia, che arriva a coprire anche alcune classi 'nucleari' di verbi inaccusativi (Cerrone & Miola 2011; Miola 2017); cfr. (24) e (25). La grammaticalizzazione di *avere* parrebbe quindi più avanzata in varietà locali che nella varietà di koinè (e in italiano). Di contro, il plurale dell'articolo indefinito nella varietà torinese è reso dalla forma invariabile del parti-

tivo senza articolo e, per lo meno in alcuni contesti, presenta un certo grado di obbligatorietà; mentre nelle varietà locali (come in italiano) ha forme che coincidono con le preposizioni articolate e alterna con l'assenza totale di marca (Miola 2017; Cerruti & Regis 2020); cfr. (26) e (27). La grammaticalizzazione dell'articolo partitivo parrebbe perciò, diversamente dal caso precedente, più avanzata nella varietà di koinè che nelle varietà locali (e in italiano).

- (24) *mi sun nà a Turin*
 io sono nato a Torino
 “io sono nato a Torino”
 (dialetto torinese, Torino; ParlaTO, www.kiparla.it)
- (25) *anti sta cà qui*
 in questa casa qui
l' ha nasciù 'l me nonò
 3SG.SBJ ha nato il mio nonno
 “in questa casa è nato mio nonno”
 (dialetto biellese, Brusnengo; Cerrone & Miola 2011: 199)
- (26) *i lese mai ed liber*
 2PL.SBJ leggete mai INDF.DET libri
 “non leggete mai dei libri”
 (dialetto torinese, Torino; ASIS, <http://asit.maldura.unipd.it>,
 Cerruti & Regis 2020: 665)
- (27) *la strà la fa*
 la strada 3SG.SBJ fa
di turnant
 INDF.DET-ART.DEF.M.PL tornanti
 “la strada fa dei tornanti”
 (dialetto biellese, Campiglia Cervo; ALEPO Q3635, Cerruti
 & Regis 2020: 667)

Il comportamento di volta in volta diverso di una stessa varietà di dialetto è inevitabilmente il prodotto di un insieme di fattori, linguistici ed extralinguistici. Fra questi, non è da trascurare il ruolo esercitato dalla (seppur generalmente scarsa, v. sopra) codificazione, che per lo meno negli usi scritti di varietà dialettali, sui quali essa ha più influenza, può contribuire ad arrestare un processo a un certo stadio. Quale sia tale stadio può poi dipendere dai caratteri stessi della codificazione, anche in relazione alle sue finalità. Basti pensare ad esempio che, nel caso dei dialetti italo-romanzi, la codificazione può eleggere a

lingua di riferimento l'italiano o, al contrario, selezionare i tratti più distanti dalla lingua tetto (appoggiandosi eventualmente a un modello esterno differente dall'italiano, v. ad es. Tosco 2011); può essere improntata alla varietà di koinè oppure a singole varietà locali; può concentrarsi su alcuni aspetti o settori del sistema linguistico, come le convenzioni ortografiche e l'inventario lessicale, tralasciandone altri; e così via. Va detto però che, in relazione al passo di grammaticalizzazione di varietà di dialetto, fattori come questi attendono ancora di essere esplorati in maniera mirata.

6. Conclusioni

In conclusione, uno sguardo variazionista, o più precisamente variazionale (§ 4.1), ai processi di grammaticalizzazione offre intanto l'opportunità di identificare varietà di lingua – anziché lingue – più e meno grammaticalizzate (§§ 3 e 5), e permette di indagare come il contatto fra varietà di una stessa lingua possa influire sul passo di grammaticalizzazione di certi processi e, più nel complesso, delle varietà considerate (§ 4). Dà modo inoltre di isolare le forze 'interne', universali, della grammaticalizzazione dalle forze 'esterne', locali, proprie di un certo diasistema, le une in linea di principio unidirezionali e le altre in grado di agire in più direzioni (§ 4.1); consentendo di cogliere il verso effettivo di un mutamento in atto in una lingua (o, più specificamente, in alcune sue varietà). Fornisce poi gli strumenti per valutare se le manifestazioni con le quali un fenomeno si presenta in varietà differenti costituiscano fasi diverse di uno stesso processo o, piuttosto, si configurino come esiti indipendenti (§§ 4.1 e 4.2).

Possibilmente, infine, la prospettiva variazionale offre indizi per fare previsioni sulla diffusione di un mutamento. Si può infatti essere portati a fare ipotesi diverse sul passo di grammaticalizzazione di un certo processo sapendo, ad esempio, che lo stadio più avanzato di quel processo si incontra esclusivamente in una varietà obsoleta, come l'italiano popolare; o, al contrario, sapendo che quel processo si trova in una fase più avanzata nel parlato colloquiale, il 'superregistro' dei cui tratti è particolarmente ricettiva la varietà neo-standard. Per previsioni, o ipotesi, come queste è tuttavia cruciale considerare l'interazione tra fattori linguistici ed extra-linguistici, che può essere responsabile della diffusione con tempi e modalità diverse di fenomeni che pure

hanno lo stesso grado e la stessa natura di marcatezza sociolinguistica (in una stessa lingua così come in lingue diverse).

Riferimenti bibliografici

- Alisova, Tatiana. 1965. Relative limitative e relative esplicative nell'italiano popolare. *Studi di Filologia Italiana* 23. 299-333.
- Amenta, Luisa. 2020. *Strutture tempo-aspettuali nel siciliano e nell'italiano regionale di Sicilia*. Palermo: Centro di studi filologici e linguistici siciliani.
- Ammon, Ulrich. 2003. On the social forces that determine what is standard in a language and on conditions of successful implementation. *Sociolinguistica* 17. 1-10.
- Arroyo, José Luis Blas & Vellón Lahoz, Javier. 2018. On the trail of grammaticalization in progress: has *el que* become a compound relative pronoun in the history of Spanish prepositional relative clauses?. *Probus* 30(1). 1-45.
- Auer, Peter. 2017a. The neo-standard of Italy and elsewhere in Europe. In Cerruti, Massimo, Crocco, Claudia & Marzo, Stefania (eds.), *Towards a new standard: Theoretical and empirical studies on the restandardization of Italian*, 365-374. Berlin-New York: Mouton de Gruyter.
- Auer, Peter. 2017b. Dialect change in Europe. Leveling and convergence. In Boberg, Charles, Nerbonne, John & Watt, Dominic (eds.), *The handbook of dialectology*, 159–176. Hoboken, NJ: Wiley-Blackwell.
- Auer, Peter, Hinskens, Frans & Kerswill, Paul (eds.). 2005. *Dialect change: Convergence and divergence in European languages*. Cambridge: Cambridge University Press.
- Ballarè, Silvia. 2020a. *Esiti del ciclo di Jespersen in area italo-romanza: grammaticalizzazione e contatto linguistico*. Bologna-Cesena: Caissa.
- Ballarè, Silvia. 2020b. L'italiano neo-standard oggi: stato dell'arte. *Italiano LinguaDue* 12(2). 469-492.
- Ballarè, Silvia, Cerruti, Massimo & Gorla, Eugenio. 2019. Variazione diastratica nel parlato di giovani: il caso delle costruzioni relative. In Moretti, Bruno, Kunz, Aline, Natale, Silvia & Krakenberger, Etna (a cura di), *Le tendenze dell'italiano contemporaneo rivisitate. Atti del LII Congresso Internazionale di Studi della Società di Linguistica Italiana (Berna, 6-8 settembre 2018)*, 75-94. Milano: Officinaventuno.

- Ballarè, Silvia & Inglese, Guglielmo (2021). The development of locative relative markers: from typology to sociolinguistics (and back). *Studies in language – Online-First Articles*. (<https://doi.org/10.1075/sl.20013.bal>) (Consultato il 23.07.2021).
- Ballarè, Silvia & Micheli, M. Silvia. 2018. Usi di *dove* nell'italiano contemporaneo: costruzioni relative e dinamiche di ristandardizzazione. *Linguistica e Filologia* 38. 29-56.
- Bernini, Giuliano & Ramat, Paolo. 1996. *Negative sentences in the languages of Europe: a typological approach*. Berlin: Mouton de Gruyter.
- Berretta, Monica. 1989. Tracce di coniugazione oggettiva in italiano. In Foresti, Fabio, Rizzi, Elena & Benedini, Paola (a cura di), *L'italiano fra le lingue romanze. Atti del XX Congresso Internazionale di Studi della SLI (Bologna, 25-27 settembre 1986)*, 125-150. Roma: Bulzoni.
- Berretta, Monica. 2002. Ordini marcati dei costituenti maggiori di frase: una rassegna. In Ead., *Temi e percorsi della linguistica. Scritti scelti. A cura di Silvia Dal Negro e Bice Mortara Garavelli*, 149-199. Vercelli: Mercurio.
- Berruto, Gaetano. 2007. Miserie e grandezze dello standard. In Molinelli, Piera (a cura di), *Standard e non standard tra scelta e norma. Atti del XXX Convegno della Società Italiana di Glottologia (Bergamo, 20-22 ottobre 2005)*, 13-42. Roma: Il Calamo.
- Berruto, Gaetano. 2009. Περὶ συντάξεως. Sintassi e variazione. In Ferrari, Angela (a cura di), *Sintassi storica e sincronica dell'italiano. Subordinazione, coordinazione, giustapposizione. Atti del X Congresso della Società Internazionale di Linguistica e Filologia Italiana (Basilea, 30 giugno-3 luglio 2008)*, vol. 1, 21-58. Firenze: Cesati.
- Berruto, Gaetano. 2012. *Sociolinguistica dell'italiano contemporaneo. Seconda edizione*. Roma: Carocci.
- Berruto, Gaetano. 2014. Esiste ancora l'italiano popolare? Una rivisitazione. In Danler, Paul & Konecny, Christine (a cura di), *Dall'architettura della lingua italiana all'architettura linguistica dell'Italia. Saggi in omaggio a Heidi Siller-Runggaldier*, 277-290. Frankfurt am Main: Lang.
- Bosson, Georg. 2003. Nominal and/or verbal marking of central actants. In Fiorentino, Giuliana (ed.), *Romance Objects. Transitivity in Romance Languages*, 17-47. Berlin-New York: Mouton de Gruyter.
- Brook, Marisa. 2011. One of those situations where a relative pronoun becomes a complementizer: A case of grammaticalization in progress... again. In Armstrong, Lisa (eds.), *Proceedings of the 2011 Annual Meeting of the Canadian Linguistic Association*, 1-7. (homes.chass.utoronto.ca/~cla-acl/actes2011/actes2011.html) (Consultato il 30.04.2021).

- Bybee, Joan. 2011. Usage-based theory and grammaticalization. In Narrog, Heiko & Heine, Bernd (eds.), *The Oxford Handbook of Grammaticalization*, 69-78. Oxford: Oxford University Press.
- Bybee, Joan L., Perkins, Revere D. & Pagliuca, William. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Carlier, Anne, De Mulder, Walter & Lamiroy, Béatrice. 2012. Introduction: The pace of grammaticalization in a typological perspective. *Folia linguistica* 46(2). 287-301.
- Cerrone, Pietro C. & Miola, Emanuele. 2011. La selezione degli ausiliari in un'area del Piemonte nordorientale. *Atti del Sodalizio Glottologico Milanese* 6. 196-207.
- Cerruti, Massimo. 2008. Condizioni e indizi di coniugazione oggettiva: I dialetti italiani settentrionali tra le lingue romanze. *Rivista Italiana di Dialettologia* 32. 13-38.
- Cerruti, Massimo. 2011a. Il concetto di variabile sociolinguistica a livello del lessico. *Studi italiani di linguistica teorica e applicata* 40(2). 211-231.
- Cerruti, Massimo. 2011b. Un fenomeno al confine tra morfosintassi e prosodia in una situazione di contatto verticale: le congiunzioni subordinanti *ben che, fin che, per che*. In Bombi, Raffaella, D'Agostino, Mari, Dal Negro, Silvia & Franceschini, Rita (a cura di), *Lingue e culture in contatto. Atti del X Congresso dell'Associazione Italiana di Linguistica Applicata (Bolzano, 18-19 febbraio 2010)*, 173-191. Perugia: Guerra.
- Cerruti, Massimo. 2017. Changes from below, changes from above. Relative constructions in contemporary Italian. In Cerruti, Massimo, Crocco, Claudia & Marzo, Stefania (eds.), *Towards a new standard: Theoretical and empirical studies on the restandardization of Italian*, 62-88. Berlin-New York: Mouton de Gruyter.
- Cerruti, Massimo. 2018. Code-switching in Italo-Romance: a variationist study of convergence in bilingual speech. *Lingue e linguaggio* 17(1). 83-101.
- Cerruti, Massimo. 2020. From dialect to standard: Facilitating and constraining factors. On some uses of the Italian negative particle *mica*. In Cerruti, Massimo & Tsiplakou, Stavroula (eds.), *Intermediate language varieties. Koinai and regional standards in Europe*, 125-148. Amsterdam-Philadelphia: John Benjamins.
- Cerruti, Massimo & Regis, Riccardo. 2020. Partitive determiners in Piedmontese: A case of language variation and change in a contact setting. *Linguistics* 58(3). 651-677.

- Cinque, Guglielmo. 1976. Mica. *Annali della Facoltà di Lettere e Filosofia dell'Università di Padova* 1. 101-112.
- Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Cornips, Leonie & Corrigan, Karen P (eds.). 2005. *Syntax and variation: Reconciling the biological and the social*. Amsterdam-Philadelphia: John Benjamins.
- Couper-Kuhlen, Elizabeth. 2011. Grammaticalization and conversation. In Narrog, Heiko & Heine, Bernd (eds.), *The Oxford Handbook of Grammaticalization*, 424-437. Oxford: Oxford University Press.
- Crocco, Claudia & Badan, Linda. 2020. Dislocazioni a destra interrogative tra grammatica e discorso. *Revue Romane – Online-First Articles*. (<https://doi.org/10.1075/rro.18017.cro>) (Consultato il 30.04.2021.)
- Croft, William. 2010. The origins of grammaticalization in the verbalization of experience. *Linguistics* 48(1). 1-48.
- De Mulder, Walter & Lamiroy, Béatrice. 2012. Gradualness of grammaticalization in Romance. The position of French, Spanish and Italian. In Breban, Tine, Vanderbiesen, Jeroen, Davidse, Kristin, Brems, Lieselotte & Mortelmans, Tanja (eds.), *Grammaticalization and Language Change*, 199-226. Amsterdam-Philadelphia: John Benjamins.
- Diewald, Gabriele. 2002. A model for relevant types of contexts in grammaticalization. In Wischer, Ilse & Diewald, Gabriele (eds.), *New Reflections on Grammaticalization*, 103-120. Amsterdam-Philadelphia: John Benjamins.
- Diewald, Gabriele. 2011. Grammaticalization and Pragmaticalization. In Narrog, Heiko & Heine, Bernd (eds.), *The Oxford Handbook of Grammaticalization*, 450-461. Oxford: Oxford University Press.
- Digesto, Salvatore. 2019. *Verum a fontibus haurire A Variationist Analysis of Subjunctive Variability Across Space and Time: from Contemporary Italian back to Latin*. Ottawa: University of Ottawa. (Tesi di dottorato.)
- Fiorentino, Giuliana. 1999. *Relativa debole. Sintassi, uso, storia in italiano*. Milano: FrancoAngeli.
- Giacalone Ramat, Anna. 2017. Grammaticalizzazione e modelli di mutamento linguistico. In Velinova, Malinka (éd.), *Normes et grammaticalisation: le cas des langues romanes*, 3-28. Sofia: CU Romanistika.
- Giacalone Ramat, Anna & Mauri, Caterina. 2012. Gradualness and pace in grammaticalization: The case of adversative connectives. *Folia Linguistica* 46(2). 483-512.

- Giacalone Ramat, Anna, Mauri, Caterina & Molinelli, Piera (eds.). 2013. *Synchrony and Diachrony. A dynamic interface*, Amsterdam-Philadelphia: John Benjamins.
- Giacalone Ramat, Anna & Sansò, Andrea. 2007. The indefinite usage of *uomo* ('man') in early Italo-Romance. Grammaticalization and areality. *Archivio glottologico italiano* 92(1). 65-111.
- Guerini, Federica. 2016. Il corpus ParVa: rilevanza per la ricerca e applicazioni didattiche. In Guerini, Federica (a cura di), *Italiano e dialetto bresciano in racconti di partigiani*, 9-38. Roma: Aracne.
- Harder, Peter & Boye, Kasper. 2011. Grammaticalization and functional linguistics. In Narrog, Heiko & Heine, Bernd (eds.), *The Oxford Handbook of Grammaticalization*, 56-68. Oxford: Oxford University Press.
- Haspelmath, Martin. 2004. On directionality in language change with particular reference to grammaticalization. In Fischer, Olga, Norde, Muriel & Perridon, Harry (eds.), *Up and down the cline. The nature of grammaticalization*, 17-44. Amsterdam-Philadelphia: John Benjamins.
- Haspelmath, Martin & Michaelis, Susanne M. 2017. Analytic and synthetic: Typological change in varieties of European languages. In Buchstaller, Isabelle & Siebenhaar, Beat (eds.), *Language Variation – European Perspectives VI: Selected papers from the 8th International Conference on Language Variation in Europe (ICLaVE 8), Leipzig 2015*, 3-22. Amsterdam-Philadelphia: John Benjamins.
- Heine, Bernd. 2002. On the role of context in grammaticalization. In Wischer, Ilse & Diewald, Gabriele (eds.), *New Reflections on Grammaticalization*, 83-101. Amsterdam-Philadelphia: John Benjamins.
- Heine, Bernd & Kuteva, Tania. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd & Kuteva, Tania. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.
- Heine, Bernd & Narrog, Heiko. 2010. Grammaticalization and linguistic analysis. In Heine, Bernd & Narrog, Heiko (eds.), *The Oxford Handbook of Linguistic Analysis*, 401-423. Oxford: Oxford University Press.
- Hopper, Paul. 1991. On some principles of grammaticization. In Traugott, Elizabeth Closs & Heine, Bernd (eds.) *Approaches to Grammaticalization*, Vol. I, pp. 17-36. Amsterdam-Philadelphia: John Benjamins.
- Hymes, Dell. 2020. The Scope of Sociolinguistics. *International Journal of the Sociology of Language* 263. 67-76.

- Janda, Richard D. 2001. Beyond “pathways” and “unidirectionality”: on the discontinuity of language transmission and the counterability of grammaticalization. *Language Sciences* 23. 265-340.
- Kortman, Bernd (ed.) 2004. *Dialectology Meets Typology: Dialect Grammar from a Cross-Linguistic Perspective*. Berlin-New York: Mouton de Gruyter.
- Kortmann, Bernd & Schneider, Agnes. 2011. Grammaticalization in Non-standard Varieties of English. In Narrog, Heiko & Heine, Bernd (eds.), *The Oxford Handbook of Grammaticalization*, 263-278. Oxford: Oxford University Press.
- Kühl, Karoline & Braunmüller, Kurt. 2014. Linguistic stability and divergence: An extended perspective on language contact. In Braunmüller, Kurt, Höder, Steffen & Kühl, Karoline (eds.), *Stability and divergence in language contact. Factors and mechanisms*, 13-38. Amsterdam-Philadelphia: John Benjamins.
- Labov, William. 2001. *Principles of Linguistic Change*. Vol. II. *Social Factors*. Oxford: Blackwell.
- Labov, William. 2007. Transmission and diffusion. *Language* 83. 344-387.
- Laitinen, Lea. 2004. Grammaticalization and standardization. In Fischer, Olga, Norde, Muriel & Perridon, Harry (eds.), *Up and down the cline. The nature of grammaticalization*, 247-262. Amsterdam-Philadelphia: John Benjamins.
- Lamiroy, Béatrice. 2011. Degrés de grammaticalisation à travers les langues de même famille. *Mémoires de la Société de linguistique de Paris* 19. 167-192.
- Lamiroy, Béatrice & De Mulder, Water. 2011. Degrees of grammaticalization across languages. In Narrog, Heiko & Heine, Bernd (eds.), *The Oxford Handbook of Grammaticalization*, 302-318. Oxford: Oxford University Press.
- Lamiroy, Béatrice & Pineda, Anna. 2017. Grammaticalization across Romance languages and the pace of language change. The position of Catalan. *Linguisticae Investigationes* 40(2). 304-331.
- Léglise, Isabelle & Chamoreau, Claudine (eds.). 2013. *The interplay of variation and change in contact settings*. Amsterdam-Philadelphia: John Benjamins.
- Lehmann, Christian. 1982. Universal and typological aspects of agreement. In Seiler, Hansjakob & Stachowiak, Franz J. (eds.), *Apprehension. Das sprachliche Erfassen von Gegenständen*. Vol. II. *Die Techniken und ihr Zusammenhang in Einzelsprachen*, 201-267. Tübingen: Narr.

- Lehmann, Christian. 1985. Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile* 20. 303-318.
- Manzotti, Emilio & Rigamonti, Alessandra. 1991. La negazione. In Renzi, Lorenzo & Salvi, Giampaolo (a cura di), *Grande grammatica italiana di consultazione*. Vol. II. *I sintagmi verbale, aggettivale, avverbiale. La subordinazione*, 245-317. Bologna: Il Mulino.
- Matras, Yaron. 2011. Grammaticalization and language contact. In Narrog, Heiko & Heine, Bernd (eds.), *The Oxford Handbook of Grammaticalization*, 279-290. Oxford: Oxford University Press.
- Meyerhoff, M. 2013. Syntactic variation and change: The variationist framework and language contact. In Légise, Isabelle & Chamoreau, Claudine (eds.), *The interplay of variation and change in contact settings*, 23-51. Amsterdam-Philadelphia: John Benjamins.
- Miller, Jim. 2004. Perfect and resultative constructions in spoken and non-standard English. In Fischer, Olga, Norde, Muriel & Perridon, Harry (eds.), *Up and down the cline. The nature of grammaticalization*, 229-246. Amsterdam-Philadelphia: John Benjamins.
- Milroy, James & Milroy, Lesley. 1985. Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21. 339-384.
- Milroy, James & Milroy, Lesley. 1997. Varieties and Variation. In Coulmas, Florian (ed.), *The Handbook of Sociolinguistics*, 47-64. Oxford: Blackwell.
- Miola, Emanuele. 2017. The position of Piedmontese on the Romance grammaticalization cline. *Folia Linguistica* 51(1). 133-167.
- Miola, Emanuele. 2018. Quali fattori accelerano il passo della grammaticalizzazione? Un'indagine su quattro varietà romanze non-standardizzate. *Archivio Glottologico Italiano* 102. 108 – 138.
- Mosegaard Hansen, Maj-Britt & Visconti, Jacqueline. 2012. The evolution of negation in French and Italian: Similarities and differences. *Folia Linguistica* 46(2). 453-482.
- Murelli, Adriano. 2011. *Relative constructions in European non-standard varieties*. Berlin-New York: Mouton de Gruyter.
- Narrog, Heiko & Otori Toshio. 2011. Grammaticalization in Japanese. In Narrog, Heiko & Heine, Bernd (eds.), *The Oxford Handbook of Grammaticalization*, 775-785. Oxford: Oxford University Press.
- Nevalainen, Terttu & Palander-Collin, Minna. 2011. Grammaticalization and sociolinguistics. In Narrog, Heiko & Heine, Bernd (eds.), *The Oxford Handbook of Grammaticalization*, 118-129. Oxford: Oxford University Press.

- Pichler, Heike (ed.). 2016. *Discourse-pragmatic variation and change in English: New methods and insights*. Cambridge: Cambridge University Press.
- Poplack, Shana. 2011. Grammaticalization and linguistic variation. In Narrog, Heiko & Heine, Bernd (eds.), *The Oxford Handbook of Grammaticalization*, 209-224. Oxford: Oxford University Press.
- Poplack, Shana, Torres Cacoullos, Rena, Dion, Nathalie, de Andrade Berlinck, Rosane, Digesto, Salvatore, Lacasse, Dora & Steuck, Jonathan. 2018. Variation and Grammaticalization in Romance: A Cross-Linguistic Study of the Subjunctive. In Ayres-Bennett, Wendy & Carruthers, Janice (eds.), *Manual of Romance Sociolinguistics*, 217-252. Berlin-New York: Mouton de Gruyter.
- Prandi, Michele. 2018. Per una razionalizzazione della grammatica e del suo insegnamento. In Fiorentino, Giuliana, Ricci, Cecilia & Siekiera, Anna (a cura di), *Trasversalità delle lingue e dell'analisi linguistica*, 19-34. Firenze: Franco Cesati.
- Ramat, Paolo. 2006. Italian negation from a typological/areal point of view. In Grandi, Nicola & Iannàccaro, Gabriele (a cura di), *Zhi. Scritti in onore di Emanuele Banfi in occasione del suo 60° compleanno*, 355-370. Cesena-Roma: Caissa.
- Regis, Riccardo. 2014. Può un dialetto essere standard?. *Vox Romanica* 72. 151-169.
- Regis, Riccardo. 2018. Formazione di varietà territoriali. In Krefeld, Thomas & Bauer, Roland (a cura di), *Lo spazio comunicativo dell'Italia e delle varietà italiane*. (www.kit.gwi.uni-muenchen.de/?p=12794&v=2) (Consultato il 30.04.2021.)
- Renzi, Lorenzo. 2019. Ancora su come cambia la lingua. Qualche nuova indicazione. In Moretti, Bruno, Kunz, Aline, Natale, Silvia & Krakenberger, Etna (a cura di), *Le tendenze dell'italiano contemporaneo rivisitate. Atti del LII Congresso Internazionale di Studi della Società di Linguistica Italiana (Berna, 6-8 settembre 2018)*, 13-33. Milano: Officinaventuno.
- Seiler, Guido. 2018. Synchrony and Diachrony – Two outdated dimensions?. In Glessgen, Martin, Kabatek, Johannes & Völker, Harald (éd.), *Repenser la variation linguistique. Actes du Colloque DIA IV à Zurich (12-14 sept. 2016)*, 77-96. Strasbourg: ELiPhi.
- Serianni, Luca. 1989. *Grammatica italiana: italiano comune e lingua letteraria. Suoni, forme, costrutti*. Torino: UTET.

- Squartini, Mario. 2017. Italian non-canonical negations as modal particles: information state, polarity and mirativity. In Sansò, Andrea & Fedriani, Chiara (eds.), *Discourse Markers, Pragmatic Markers and Modal Particles: New Perspectives*, 203-229. Amsterdam-Philadelphia: John Benjamins.
- Tagliamonte, Sali A. 2000. The grammaticalisation of the present perfect in English: Tracks of change and continuity in a linguistic enclave. In Fischer, Olga, Rosenbach, Anette & Stein, Dieter (eds.), *Pathways of Change: Grammaticalization in English*, 329-354. Amsterdam-Philadelphia: John Benjamins.
- Torres Cacoullous, Rena. 2011. Variation and grammaticalization. In Díaz-Campos, Manuel (ed.), *The Handbook of Hispanic Sociolinguistics*, 148-167. Oxford: Blackwell.
- Tosco, Mauro. 2011. Between endangerment and Ausbau. In Miola, Emanuele & Ramat, Paolo (eds.), *Language contact and language decay: Socio-political and linguistic perspectives*, 227-46. Pavia: IUSS Press.
- Traugott, Elizabeth Closs. 2011. Grammaticalization and mechanisms of change. In Narrog, Heiko & Heine, Bernd (eds.), *The Oxford Handbook of Grammaticalization*, 19-30. Oxford: Oxford University Press.
- Traugott, Elizabeth Closs & Trousdale, Graeme 2010. Gradience, gradualness and grammaticalization: how do they intersect?. In Traugott, Elizabeth Closs & Trousdale, Graeme (eds.), *Gradience, Gradualness and Grammaticalization*, 19-44. Amsterdam-Philadelphia: John Benjamins.
- Vismans, Roel, Hüning, Matthias & Weerman, Fred (eds.). 2010. *Dutch between English and German*. Special issue of *Journal of Germanic Linguistics* 22(4).

NICOLA GRANDI

Fattori sociolinguistici e costruzione del campione tipologico. Su alcune possibili interazioni tra tipologia e sociolinguistica¹

Una delle distorsioni che occorre evitare nella costruzione di un campione tipologico è legata al numero dei parlanti di una lingua. È ovvio che il successo o l'insuccesso di una lingua non dipendono da alcun tratto strutturale. Nella grammatica di una lingua, cioè, non c'è nulla che possa orientarne i destini. Molto meno ovvio è che il numero dei parlanti non possa influenzare la struttura del sistema di una lingua e che, quindi, non possa essere considerato come un parametro pertinente per un'indagine tipologica. In questo contributo cercherò di mostrare come la dimensione e l'articolazione interna delle comunità linguistiche possano in effetti condizionare la configurazione tipologica, focalizzando tre situazioni. La prima riguarda il rapporto tra diversità inter- e intralinguistica. La seconda concerne la tendenza dei tipi più rari o 'recessivi' a stabilizzarsi in comunità ridotte e isolate. La seconda si riferisce invece alle possibili peculiarità della varietà standard-normativa all'interno di diasistemi complessi in uso in comunità molto numerose e internamente articolate.

Parole chiave: campione rappresentativo, tipi recessivi, varietà di lingua, isolamento, diasistema.

1. Introduzione

Peter Trudgill (2011: viii) afferma che “different types of human society produce different types of language”. Quindi l'estrema varietà di tipi di lingue che è possibile osservare oggi al mondo sarebbe anche frutto dei diversi modelli di società che caratterizzano le comunità umane². Questa

¹ Questo contributo presenta alcuni risultati di una indagine sul rapporto tra tipologia e sociolinguistica condotta nell'ambito del progetto PRIN 2017 AH4 2017LAP429 *UniverS-Ita L'italiano scritto degli studenti universitari: quadro sociolinguistico, tendenze tipologiche, implicazioni didattiche*.

² Sul tema segnalo anche il recente Sinnemäki & Di Garbo (2018).

considerazione pone una domanda di ricerca di grande interesse potenziale: la variazione interlinguistica può avere a che fare con la struttura delle società ed anche con il contesto geografico nel quale queste si sono sviluppate?

Se questo fosse vero, la tipologia linguistica, che programmaticamente ambisce a indagare i limiti e le regolarità della variazione linguistica, non dovrebbe prescindere da una interazione sistematica con la sociolinguistica e dalla ricerca di correlazioni non solo tra strutture linguistiche ricorrenti e principi di natura funzionale, ma anche tra strutture linguistiche ricorrenti e contesti e configurazioni sociali.

L'ipotesi su cui lavorare, dunque, è che le configurazioni socio-ambientali possano favorire o sfavorire alcune specifiche configurazioni 'grammaticali' o, meglio, 'sistemiche'. In quest'ottica, alcuni possibili parametri cui fare riferimento per interpretare i dati linguistici sarebbero la consistenza demografica delle comunità linguistiche, la loro collocazione geo-fisica, il modello di sviluppo socioculturale-economico-politico che le contraddistingue, la presenza e le caratteristiche dei sistemi di istruzione, ecc.

In questo contributo vorrei indagare alcune possibili conseguenze di questa domanda di ricerca, con riferimento principalmente alla costruzione di un campione per una indagine tipologica e al grado di rappresentatività delle lingue corrispondenti a comunità umane più e meno numerose.

2. Le possibili distorsioni nella costruzione del campione: il numero dei parlanti

In tipologia, tradizionalmente, le lingue sono 'astratte' dal loro divenire storico³. Ciò significa sia che le lingue sono 'fotografate' nella loro configurazione strutturale indipendentemente dai processi di mutamento linguistico che hanno prodotto tale configurazione, sia che questa configurazione è considerata a prescindere dai fattori storici e geografici che hanno condizionato le vicende della comunità linguistica.

Una delle conseguenze più immediate di tale assunto è che nella costruzione di un campione rappresentativo per un'indagine tipolo-

³ Va detto, tuttavia, che gli approcci più recenti, di tipo *source-oriented*, si fondano su una integrazione più stabile tra prospettiva diacronica e sincronica. Si veda, tra gli altri, Cristofaro (2019).

gica si raccomanda sempre di non tener conto del numero di parlanti di una lingua⁴. La motivazione più immediata è banale: le lingue più parlate sono di norma anche quelle meglio descritte, per le quali, dunque, è più facile reperire dati, sia attraverso questionari, sia attraverso grammatiche; spesso, quindi, esse vengono selezionate in modo quasi inerziale. Tuttavia, dietro l'indicazione di non considerare il numero di parlanti nella scelta delle lingue del campione c'è una ragione più profonda: il numero di parlanti di una lingua non dipende mai da criteri interni alla lingua stessa, ma solo dai successi o dagli insuccessi di chi la parla. In altri termini, non ci sono specifiche configurazioni grammaticali che possono favorire o sfavorire una lingua. E neppure la maggiore coerenza tipologica gioca un ruolo nell'avvantaggiare alcuni idiomi rispetto ad altri. Se così fosse, l'inglese, lingua massimamente incoerente sia sul piano morfologico che su quello sintattico, dovrebbe essere oggi relegata al rango di lingua minoritaria.

Tuttavia se osservare che nulla, in una grammatica, può farci prevedere se una lingua avrà o meno successo (quindi più o meno parlanti) è assolutamente indubbio, assai meno scontato è asserire il contrario, cioè che il successo di una lingua non ci consente di operare previsioni sulla sua struttura.

Questo aspetto, spesso trascurato nella letteratura tipologica, merita, a mio modo di vedere, una certa attenzione. Vorrei dunque affrontarlo focalizzando tre casi di studio: il rapporto tra variazione inter- e intralinguistica e il ruolo che quest'ultima può giocare in una indagine tipologica; la conservazione di tratti o tipi 'recessivi' nelle lingue 'piccole' e, infine, le possibili peculiarità delle varietà fortemente standardizzate nei diasistemi più complessi.

2.1 Variazione inter- e intralinguistica

In tipologia, il dibattito sulla diversità / variazione è totalmente 'appiattito' sulla diversità / variazione interlinguistica. La tipologia, cioè, studia, programmaticamente, lingue diverse. In questo senso, è notissima la correlazione tra i tassi di diversità interlinguistica e i tassi di biodiversità: le zone del pianeta in cui si concentra il maggior numero di specie del regno animale e vegetale sono anche quelle in cui la diversità tra lingue

⁴ Sul tema si consideri, tra gli altri, Song (2018: 88-91).

si mostra più rigogliosa. Al proposito, Nettle & Romaine (2001: 28) hanno introdotto l'etichetta di "diversità biolinguistica".

Se si osservasse una mappa della distribuzione delle lingue del mondo, si osserverebbero zone con addensamenti molto fitti, concentrate principalmente a cavallo dell'Equatore: in esse, sono presenti moltissime lingue cui corrispondono comunità scarsamente popolate. E zone, come l'Europa settentrionale, centrale e occidentale, in cui la presenza di lingue pare diradarsi. Si tratta però di regioni molto popolate, in cui vi sono sì poche lingue, ma con comunità di riferimento molto numerose. In altri termini, si configurano due situazioni opposte: regioni con molte lingue con pochi parlanti e regioni con poche lingue con molti parlanti. Le prime sono sovente considerate, per la tipologia, un terreno di indagine assai più stimolante delle seconde.

Gnerre (2011: 132) traccia un quadro delle variabili che concorrono a condizionare e indirizzare i tassi di diversità linguistica:

- i) le dimensioni demografiche delle comunità linguistiche
- ii) le stratificazioni socio-economiche o religiose al loro interno
- iii) la presenza o assenza di una varietà linguistica di prestigio o standardizzata rispetto alle altre
- iv) il controllo socio-politico (interno o esterno) esercitato su di esse
- v) l'assenza (nella maggior parte dei casi) o la presenza (in una minoranza di casi, e solo nel corso degli ultimi millenni) di qualche forma di scrittura
- vi) l'uso che di questa, se presente, viene fatto.

Ciascuna variabile è indipendente dalle altre e può essere idealmente considerata un *continuum* tra due poli contrapposti, positivo e negativo. Valori orientati verso il polo negativo tendono a delineare una situazione di estrema frammentazione della presenza umana con livelli alti di diversità linguistico-culturale (quindi molte lingue diverse con gruppi sociali retrostanti mediamente ridotti). Invece, valori orientati verso il polo positivo tendono a definire uno scenario contraddistinto da una distribuzione più omogenea e continua della presenza umana, con comunità numericamente consistenti e una riduzione dei tassi di diversità linguistico-culturale. In questo caso, tuttavia, aumenta esponenzialmente il tasso di diversità intralinguistica (pressoché assente nella situazione precedente): i diasistemi si fanno più complessi e articolati e la forbice che distingue le varietà diatopiche, diastratiche e diafasiche di una lingua si divarica. Possiamo considerare la prima

diversità come ‘orizzontale’, cioè realizzata principalmente come una successione di lingue diverse nello spazio fisico; e la seconda come ‘verticale’, cioè realizzata principalmente come una successione di varietà diverse negli strati di una società/comunità.

Se dunque la diversità è osservata da un punto di vista che contempli sia la prospettiva ‘orizzontale’ che quella ‘verticale’, le variabili citate sopra a partire da Gnerre (2011) mostrano un potenziale molto più ampio, consentendo di fare previsioni non solo sulla quantità di lingue diverse, ma anche sulla natura dei diasistemi, quindi sulla quantità di varietà diverse, e giocano dunque un ruolo cruciale nell’elaborazione di una teoria complessiva della diversità. Diversità inter- e intralinguistica tendono ad avere, perciò, un rapporto di proporzionalità inversa (Grandi 2020: 422).

Questo pone per la tipologia e, più in generale, per una teoria complessiva della diversità linguistica un problema concettuale e, poi, metodologico di rilievo. Si considerino, ad esempio, i dati seguenti, che esemplificano il fenomeno del cosiddetto accusativo preposizionale:

(1)⁵ Spagnolo

<i>Los</i>	<i>ácidos</i>	<i>atacan</i>	<i>a</i>
ART.DEF.M.PL	acids	attack.3PL	ACC
<i>los</i>	<i>metales</i>		
ART.DEF.M.PL	metals		
‘Acids attack metals’			
(Brugè & Brugger 1994)			

(2) Rumeno

<i>Toti</i>	<i>bărbații</i>	<i>(o)</i>	<i>iubesc</i>	<i>pe</i>
All	men	(her)	love	ACC
<i>o</i>		<i>femeie</i>		
ART.INDF.F.SG		woman		
‘All men love a woman’ [+ specific]				
(Cornilescu 2000)				

In italiano standard, nelle frasi non marcate il complemento oggetto non è introdotto da una preposizione, ma è realizzato in modo ‘posizionale’, cioè è collocato dopo il verbo (es. *ho incontrato mio padre*); l’oggetto preposizionale è però largamente utilizzato in alcune forme

⁵ Nelle glosse e nella traduzione dei dati è stata mantenuta la lingua della fonte da cui sono tratti.

non standard marcate in diatopia (ad es. in varietà meridionali e insulari) e in diastratia e/o diafasia (ad es. nel parlato non sorvegliato: *e a te chi ti ha chiamato?, vai a chiamare a tuo fratello*; cfr. Fiorentino 2010). Si tratta di varietà che la tipologia generalmente trascura, in quanto poco rappresentate nelle grammatiche di riferimento dell'italiano. Tuttavia, la differenza / distanza tipologica che intercorre tra italiano standard *ho incontrato tuo padre* e italiano substandard regionale *ho incontrato a tuo padre* è la stessa che c'è tra italiano *ho conosciuto tutti i tuoi amici* e spagnolo *he conocido a todos tus amigos*.

“Questo rende evidente un assunto a lungo trascurato in letteratura: a livello empirico e conoscitivo la dimensione in cui si manifesta la diversità linguistica è irrilevante. Se due strutture sono diverse, lo sono a prescindere dall'essere riconosciute come peculiari di lingue differenti o di varietà del medesimo diasistema. Ciò rende la diversità verticale un terreno di indagine molto promettente per la tipologia esattamente quanto la diversità orizzontale e dovrebbe indurci a rassegnarci all'idea della convivenza, in un diasistema, di strutture tipologicamente anche molto distanti” (Grandi 2020: 427).

Le varietà di lingua hanno perciò, rispetto all'indagine tipologica, la stessa legittimità che viene attribuita alle lingue diverse, in quanto identico è il modo in cui la diversità si manifesta sul piano strutturale. In questo quadro, il numero di parlanti è una variabile che la tipologia non dovrebbe accantonare, in quanto è proprio il numero di parlanti, unitamente al grado di articolazione interna della comunità, a orientare il piano su cui questa diversità tende a manifestarsi: quello 'orizzontale' nel caso di lingue 'piccole'; quello 'verticale' nel caso di lingue 'grandi'.

Queste considerazioni pongono, come si è detto poco sopra, alcuni problemi metodologici, soprattutto rispetto alla rappresentatività e al bilanciamento del campione. Si consideri, ad esempio, il caso di lingue, come l'italiano, cui corrispondono comunità con demografie molto numerose e con reti sociali dense o addirittura 'rinforzate' da strumenti supplementari di interazione come il telefono o internet e caratterizzate da forti stratificazioni socio-economiche (Gnerre 2011: 133). Lingue di questo tipo prevedono di norma diasistemi complessi, nei quali in genere è presente una varietà linguistica standardizzata e considerata più prestigiosa delle altre. In questo quadro, che fare, in una indagine tipologica, delle varietà diatopiche / distriche / diafa-

siche? Come scegliere la varietà da inserire nel campione in rappresentanza dell'intero diasistema, posto che sia davvero necessario scegliere sempre solo una varietà per ogni diasistema? Cosa cambierebbe in ciò che conosciamo della variazione interlinguistica se scegliessimo sempre la varietà statisticamente più usata e non, come avviene pressoché di *default*, quella standard e normativa?

Si tratta di domande cruciali per la tipologia e per l'affidabilità dei risultati delle indagini comparative su larga scala. Come ha notato Berretta (2002[1994]), negli studi tipologici non si considera che per le lingue fortemente diversificate in diatopia, diafasia e diastratia la scelta della varietà di lingua da assumere a riferimento può condizionare in modo decisivo il risultato dell'indagine.

Se, per riprendere i dati citati sopra, l'indagine tipologica fosse focalizzata sulla codifica dell'oggetto diretto, l'italiano potrebbe essere descritto sia come lingua che prevede una sua realizzazione esclusivamente 'posizionale', cioè attraverso l'ordine dei costituenti, sia come lingua che alterna una codifica 'posizionale' e una codifica analitica, mediante l'accusativo preposizionale. Tutto dipende, come è ovvio, da cosa si osserva e da quale varietà viene assunta a riferimento. Nel primo caso la generalizzazione copre la varietà standard normativa e molte varietà del Nord, ma esclude quelle del Sud; nel secondo caso, la generalizzazione copre soprattutto le varietà meridionali e insulari, ma non lo standard. Ne consegue che nessuna delle due generalizzazioni rappresenta fedelmente l'intero diasistema; e, quindi, che ogni scelta tra le varietà limita la prospettiva e restituisce un punto di osservazione sulla lingua parziale e dunque limitatamente affidabile. Ma ciò pone un interrogativo ancor più stringente: possono esserci, in un campione, lingue più affidabili di altre?

Intuitivamente si potrebbe affermare che la soluzione a questi problemi sia quella di privilegiare, nella costruzione del campione e quindi nell'indagine tipologica, lingue che non pongano questo problema, cioè lingue con comunità di parlanti demograficamente ridotte e con un tasso limitato di variazione interna. È di questa opinione Dik Bakker (2012: 6): "if population size factor may indeed turn out to be fundamental, then a sample [...] should consistently contain a relative overrepresentation of the smallest languages".

Questa scelta parrebbe legittima, se si considera che in effetti le dinamiche linguistiche tipiche dei diasistemi complessi sono spesso

un prodotto socio-storico relativamente recente, per molti versi legato a politiche di colonizzazione e/o standardizzazione che non sono troppo rappresentative della storia umana a lungo termine. Tuttavia, anche le ‘lingue piccole’ non sono immuni da problemi di rappresentatività, che illustrerò nella prossima sezione, prima di tornare nuovamente ai diasistemi più articolati.

2.2 Lingue ‘piccole’ e tipi ‘recessivi’

Daniel Nettle (1999: 139) osserva che “in population genetics, it is well known that the effects of random drift are greater when the population is small. This is because the probability of a slightly deleterious variant becoming fixed in a population is inversely related to the population size [...]. The smaller the community, the greater the stochastic effects of chance changes in gene frequency”. In sostanza, gli effetti della deriva genetica (ad esempio nella riduzione di variabilità, nello sbilanciamento di alcune varianti su altre, nella permanenza di alleli recessivi, ecc.) sono proporzionali alle dimensioni delle popolazioni. Nelle comunità piccole, dunque, la probabilità che un allele recessivo si stabilizzi è più alta, principalmente per effetto del maggior numero di accoppiamenti tra consanguinei. Questa probabilità aumenta se le comunità piccole sono anche isolate. Invece, nelle comunità più popolose gli alleli recessivi tendono ad essere, per così dire, riassorbiti nel quadro di una più ampia variabilità. “In general, the smaller the community, the greater the probability that a given variant that has no functional advantage at all, but is neutral or slightly disadvantageous, can replace the existing item and become the norm” (Nettle 1999: 139). L’aspetto delicato è capire come questa considerazione possa essere proiettata sulla distribuzione dei tipi linguistici. Come è noto, alcuni tipi sono molto più diffusi di altri. In alcuni casi, anche in modo statisticamente schiacciante, come avviene, ad esempio, nel caso dell’ordine dei costituenti nella frase dichiarativa assertiva, che vede le configurazioni in cui il soggetto precede l’oggetto diretto attestate nell’84% delle lingue del mondo, mentre i tipi in cui l’oggetto diretto è anteposto al soggetto occorrono solo nel 3% delle lingue⁶. E possono, dunque, essere in un certo senso definiti ‘recessivi’. La distri-

⁶ Nel restante 13% non c’è un ordine prevalente (Dryer 2013). Il WALS è stato consultato, per tutti i dati riportati in questo contributo, il 5 settembre 2020.

buzione dei tipi legati alla struttura della frase indipendente dichiarativa assertiva rivela in modo inequivocabile un qualche vantaggio funzionale delle strutture in cui il soggetto precede l'oggetto diretto⁷. La questione è capire se esista una correlazione tra tipi devianti rispetto a questo *pattern* e demografie ridotte ed eventualmente isolate, alla stregua di quanto osservato sopra per la deriva genetica. In altri termini, ci si può domandare se sia lecito parlare, per la distribuzione dei tipi 'recessivi', di una sorta di 'effetto del fondatore', non tanto nel senso di un piccolo gruppo di 'pionieri' che si staccano da una comunità, si isolano da essa e colonizzano un nuovo ambiente, quando piuttosto nel senso della perdita di variabilità, attraverso la scomparsa di alcune linee genetiche (= configurazioni tipologiche) e la conseguente maggior incidenza di alleli recessivi e di malattie genetiche (= tratti tipologici rari).

Nettle (1999: 139) ipotizza che

the rare, non-optimal orders would be more likely to be found in small communities than in large ones, since these would be more vulnerable to drift away from optimal states.

Per testare questa ipotesi, al netto delle difficoltà legate al computo esatto di parlanti di lingue spesso scarsamente documentate, Nettle (1999: 140) analizza 19 lingue con oggetto iniziale: la media dei parlanti è 750. La media dei parlanti delle lingue del mondo è di circa 5.000 (Nettle la calcola sui dati del database SIL):

it does thus seem that there is an association between object-initial order and community size [...]. [I]t seems that the aberrant, object-initial word orders may indeed be more likely to be found in small communities (Nettle 1999: 140-141)⁸.

⁷ Su questo tema e sulla distribuzione di pattern diversi rispetto all'ordine dei costituenti lavori basati su dati quantitativi possono consentire un significativo avanzamento delle conoscenze. In questa sede cito, tra gli altri, il recente Levshina (2019).

⁸ Uno dei revisori anonimi della collana mi segnala, a mio parere molto opportunamente, che le lingue 'piccole' sono assai più numerose delle lingue 'grandi': più o meno in un rapporto di 7 a 1. Ne consegue che i tratti 'recessivi' sono ovviamente più frequenti in lingue 'piccole' in quanto la possibilità che una lingua sia 'grande' e con un tratto 'recessivo' è decisamente inferiore e quella che una lingua sia 'piccola' e con un tratto 'recessivo'. Io credo, in effetti, che nella persistenza dei tratti recessivi conti soprattutto il combinato della ridotta demografia e dell'isolamento, più che il primo parametro in sé (su questo aspetto tornerò a breve). In effetti, le lingue nella

Possiamo chiederci se questa tendenza venga ulteriormente esasperata qualora al parametro della demografia ridotta si aggiunga quello dell'isolamento estremo. In Grandi (2013) ho passato in rassegna 24 lingue insulari, analizzandole per 183 tratti tipologici censiti sul WALS. Rispetto all'ordine dei costituenti nella frase indipendente dichiarativa assertiva il quadro sembra suggerire che l'isolamento accentui nettamente la tendenza appena osservata. Nelle lingue insulari censite, le lingue con ordine SOV e SVO sono il 37,4% (rispetto al 76,7% osservato nell'intero campione del WALS); le lingue insulari con oggetto iniziale sono l'8,3% (rispetto al 3% scarso nel campione complessivo del WALS). Occorre precisare che buona parte di queste lingue appartiene alla famiglia austronesiana e questo, ovviamente, 'sbilancia' il campione dal punto di vista areale e genealogico. Tuttavia, questo dato pare confermare la considerazione appena svolta. Secondo Banfi (2008: 267) l'ordine del protoaustronesiano era VSO; questo ordine è poi evoluto o verso lo schema SVO (prevalente ed attestato, ad esempio, tra le lingue malayo-polinesiane e melanesiane occidentali e orientali) o verso pattern a verbo iniziale, che si concretizzano sia nel tipo VSO che nel tipo VOS⁹, entrambi piuttosto inusuali tipologicamente, seppur con gradi diversi¹⁰. Entrambi, ma il secondo in modo particolare, paiono attestati soprattutto in lingue insulari e isolate, come il kiribati o il cèmuhi.

Quindi l'isolamento pare giocare un ruolo cruciale nel favorire la tendenza a conservare tratti tipologici 'recessivi' osservata per le lingue 'piccole'. Per altro, è possibile stabilire una sorta di nesso causale, seppur non biunivoco, tra queste due caratteristiche, in quanto se è vero che non tutte le lingue 'piccole' sono isolate, è assai frequente che le lingue isolate siano anche 'piccole' (a maggior ragione se le lingue sono isolate in quanto... isolate).

Ciò induce a considerare un ulteriore parametro che può incidere sulla configurazione tipologica delle lingue 'piccole'. Trudgill (2011:

lista di Nettle (1999: 140) sono tutte in condizioni di isolamento, più o meno marcato. Resta comunque un punto cruciale, a supporto dell'argomentazione di Nettle (1999): nelle lingue 'grandi' i tipi recessivi hanno una tendenza maggiore ad essere 'riassorbiti'.

⁹ Cfr. Himmelmann (2005) e Polinsky & Potsdam (in stampa).

¹⁰ Il tipo VOS deve la sua 'rarietà' all'anteposizione dell'oggetto diretto rispetto al verbo. Il tipo VSO, più frequente del precedente, mostra invece una particolare disposizione di tema e rema, con il tema (tipicamente il soggetto) che 'interrompe' il rema (tipicamente il sintagma verbale).

62-90) individua infatti una correlazione tra isolamento e processi di ‘complicazione’, che si concretizzano, ad esempio, nell’aumento di opacità, di ridondanza sintagmatica e paradigmatica, di irregolarità e nell’aggiunta di categorie morfologiche. Le condizioni sociali che paiono favorire questi processi sono in effetti quelle che determinano un’esposizione bassa, o nulla, al contatto: “it is in low-contact communities that we are most likely to find not only the preservation of complexity but also an *increase* in complexity” (2011: 64)¹¹. In termini generali, dunque, Trudgill (2011: 72) osserva che “there is a tendency for complexification to more common – probably a great deal more common – in low-contact than in high-contact varieties”.

Il quadro che si delinea, dunque, è quello in cui le lingue con demografie ridotte o ridottissime e isolate mostrano una propensione più alta della media a stabilizzare tipi ‘recessivi’ e con un grado di complessità tendenzialmente superiore a quello di lingue o di varietà di lingua maggiormente esposte a contatti e contaminazioni (come invece accade più spesso in comunità più numerose).

Ci si può chiedere, dunque, quanto sia ‘naturale’ questo processo e quanto tale scenario sia rappresentativo della condizione di vita ‘normale’ di una lingua¹². Sopra si è osservato come le dinamiche (socio) linguistiche innescate da comunità incredibilmente popolose e legate a processi di standardizzazione, tipiche dell’idea di Stato-Nazione sviluppatesi in epoca moderna nel mondo occidentale, siano appunto un prodotto piuttosto recente e, quindi, poco rappresentativo dei percorsi evolutivi e del comportamento delle lingue nella storia dell’uomo. E come le lingue ‘grandi’ pongano seri problemi metodologici per la costruzione di un campione tipologico.

Però, anche le peculiarità dei processi di sviluppo delle lingue ‘piccole’, soprattutto se, come spesso accade, esse sono anche isolate, esibiscono tratti di ‘innaturalità’ che ci fanno dubitare della loro capacità

¹¹ Corsivo nell’originale.

¹² Qui uso ‘normale’ in senso intuitivo, senza cioè darne una definizione esplicita. L’idea di fondo è che le due situazioni descritte nelle pagine precedenti, cioè quella di lingue molto ‘piccole’ e molto isolate e quella di lingue molto ‘grandi’ caratterizzate da processi di standardizzazione fortemente condizionati o indirizzati da interventi di tipo politico, rappresentino, seppur per ragioni diverse, casi piuttosto ‘innaturali’, rispetto alle condizioni di vita ‘medie’ delle comunità umane, in quanto entrambe subiscono una sorta di ‘forzatura’ che ha origine ambientale nel caso dell’isolamento estremo e socio-politica nel caso dei processi di standardizzazione.

di rappresentare in modo attendibile ed adeguato le ‘condizioni di vita normali’ delle lingue umane. In questo senso, la scelta di privilegiare, nella costruzione del campione, le lingue ‘piccole’, per superare i problemi di scelta tra varietà posti dai diasistemi delle lingue ‘grandi’ avrebbe, come effetto, quello di dare a tipi ‘recessivi’ e a configurazioni ‘complesse’ un peso maggiore di quello che essi concretamente hanno nella realtà¹³.

3. *Varietà standard e lingue ‘piccole’: la strana coppia*

In base a quanto affermato in precedenza, c’è un elemento che pare accomunare i diasistemi complessi e le lingue ‘piccole’: le dinamiche che caratterizzano tanto i primi, quanto le seconde paiono riprodurre situazioni sociolinguisticamente con un grado limitato di ‘naturalzza’.

Come si è detto sopra, nel caso di lingue ‘grandi’, che pongono al tipologo un problema di scelta tra varietà differenti, la varietà standard-normativa, ove presente, rappresenta in genere la scelta di *default*. Vale la pena, dunque, soffermarsi più diffusamente su essa. Per altro, proprio i processi di standardizzazione rappresentano forse l’aspetto meno naturale nelle dinamiche che contraddistinguono la storia dei diasistemi complessi, in quanto rispondono a esigenze politiche, sociali ed economiche emerse in epoche relativamente recenti. Alla luce di ciò, possiamo chiederci se i processi linguistici che si osservano nelle varietà fortemente standardizzate esibiscano delle peculiarità particolari, ad esempio conservando configurazioni tipologiche particolari e ‘bizzarre’.

Per rispondere a questa domanda sarebbe necessaria una ricognizione ampia dei *pattern* tipologici che occorrono nelle varietà standard-normative e in quelle sub-standard dei diasistemi delle ‘grandi’ lingue parlate (e scritte) nel mondo occidentale. In questa sede mi limito ad alcune osservazioni relative al diasistema italiano.

¹³ In questa affermazione e, in generale, in quanto discusso fino ad ora viene trascurato, per ragioni di spazio, un aspetto che in realtà andrebbe tenuto in conto, legato alla natura del campione tipologico che normalmente può essere, come afferma Song (2018: 83) o un *variety sample* o un *probability sample*. La scelta del campione, per altro, dipende strettamente dagli obiettivi dell’indagine. Per il caso discusso in questa sede, si può affermare che una sovrarappresentazione delle lingue ‘piccole’, ad esempio, ha senza dubbio un impatto maggiore su un *probability sample* che su un *variety sample*, per il quale può addirittura essere un elemento positivo.

In italiano, come noto, è in corso un processo di ristandardizzazione che porta all'affermazione, nell'uso pubblico, di costruzioni e strutture precedentemente confinate all'uso non sorvegliato, spesso stigmatizzare ed esplicitamente sanzionate¹⁴. È altrettanto noto come la varietà standard-normativa, imposta seguendo dinamiche di tipo *top down* sulla base di un modello colto-letterario, quindi d'*élite*, e fortemente marcato in diatopia su base toscana, sia stata e sia ancora una lingua con un numero assai limitato di parlanti nativi.

Una prima ricognizione dei dati, centrata sui segmenti del sistema per i quali si osserva una certa divaricazione tra quanto la norma prescrive e quanto i parlanti effettivamente fanno nell'uso quotidiano, rivela in effetti come la varietà standard-normativa esibisca *pattern* tipologici talvolta piuttosto inusuali.

Si consideri, ad esempio, il sistema dei dimostrativi che nella varietà normata è (meglio, era) del tipo *three way contrast*, cioè con tre elementi finalizzati a codificare ciò che è vicino a chi parla e a chi ascolta (*questo*), ciò che è lontano da chi parla, ma vicino a chi ascolta (*codesto*) e ciò che è lontano sia da chi parla, sia da chi ascolta (*quello*). Si tratta, quindi, di un sistema con due 'centri deittici', identificati nei due partecipanti alla situazione comunicativa. Un sistema a tre membri è attestato, nel campione del WALS, in 88 lingue, il 38% delle quali ha meno di 5.000 parlanti¹⁵. La quasi totalità di queste lingue è concentrata nella fascia immediatamente a nord e a sud dell'Equatore, dove si registrano i più alti tassi di diversità linguistica. Inoltre, poco meno di metà di queste lingue si trova in condizioni di isolamento talvolta estremo (spesso di tipo insulare). Tornando all'italiano, nella lingua d'uso, oggi, prevale nettamente un sistema del tipo *two way contrast*, che identifica un solo 'centro deittico' nel mittente e che codifica solo ciò che, rispetto ad egli, è vicino (*questo*) o lontano (*quello*). Le lingue che seguono questo modello, nel WALS, sono 126; la percentuale di lingue 'piccole', con meno di 5.000 parlanti, scende al 26%. In generale, le lingue di questo tipo hanno una distribuzione più diffusa e non preferenzialmente associata alla macroregione equatoriale; diminuiscono anche le situazioni di radicale isolamento¹⁶.

¹⁴ Per un quadro generale si vedano Berruto (2012) e Ballarè (2020).

¹⁵ Assumo a riferimento questo valore soglia che, come si è detto sopra, rappresenta la media dei parlanti per ogni lingua del mondo secondo la stima di Nettle (1999).

¹⁶ Diessel (2013).

Si consideri, ora, il sistema verbale del passato, che nella varietà standard dell'italiano è articolato in un complesso paradigma basato su gradi diversi di *remoteness*: passato prossimo, trapassato prossimo, passato remoto, trapassato remoto, cui si aggiunge l'imperfetto che trasmette, invece, un'informazione di tipo aspettuale. Le lingue che, nel campione del WALs, si caratterizzano per sistemi con quattro o più gradi di *remoteness* sono solo due (parlate in Bolivia e Perù) ed entrambe hanno meno di 5.000 parlanti. Quelle che esibiscono due o tre gradi di *remoteness* sono 38: il 48% di esse ha meno di 5.000 parlanti. Anche in questo caso, le lingue sono concentrate nella zona a nord e a sud dell'Equatore¹⁷.

Nella varietà d'uso dell'italiano, è invece chiaramente prevalente un sistema in cui il passato remoto, in declino in quanto fortemente marcato in diatopia e legato in modo preferenziale alle varietà meridionali, ed il passato prossimo sono nella sostanza equivalenti ed esprimono un'azione del passato con valore aspettuale perfettivo senza alcuno sfasamento di *remoteness*, in contrapposizione all'imperfetto, con valore aspettuale, appunto, imperfettivo. I vari trapassati hanno un indice di occorrenza reale in media piuttosto basso. Possiamo, dunque, descrivere questo sistema come caratterizzato da *no remoteness distinctions*: l'opposizione cardine è quella tra presente e passato, cui si aggiungono appunto informazioni di natura aspettuale. Nel WALs le lingue di questo tipo sono 94 e solo il 13% di esse ha meno di 5.000 parlanti. Inoltre, queste lingue sono distribuite in modo molto omogeneo nei diversi continenti.

In entrambi i casi (sistema dei dimostrativi e del passato) pare quindi che sopra una certa soglia di parlanti alcuni *pattern* tipologici si diradino. Nel diasistema italiano i tipi associati alla varietà standard-normativa sono proprio quelli che mostrano una frequenza maggiore nelle lingue 'piccole' e ciò induce a ipotizzare che le varietà fortemente standardizzate mostrino dinamiche non dissimili da quelle osservate in lingue con demografie estremamente ridotte e collocate in contesti di isolamento spesso estremo. Ciò che unisce due situazioni all'apparenza così diverse è una certa 'innaturalità' dei processi storici e sociolinguistici che fanno da sfondo a questi *pattern* tipologici rari, dovuta ad una sorta di 'forzatura' che agisce o per ragioni geografiche (nel caso dell'isolamento) o per ragioni socio-politiche (nel caso della standardizzazione).

Come si è visto sopra, un ulteriore aspetto che contraddistingue le lingue 'piccole', in particolare se isolate, è la conservazione o addirittura

¹⁷ Dahl & Velupillai (2013).

ra l'incremento dei tassi di complessità che si concretizza, ad esempio, nell'elevata incidenza della ridondanza sintagmatica, nel mantenimento di categorie morfologiche, nella sopravvivenza di forme opache, ecc. Possiamo chiederci cosa accada, rispetto a questo parametro, nelle varietà standardizzate. Se il quadro tracciato sopra è plausibile, esse dovrebbero essere più inclini a conservare, se non addirittura a incrementare, strutture e costruzioni complesse. E, di converso, fenomeni di semplificazione¹⁸ dovrebbero essere più frequenti nelle varietà non standard.

Molti dei fenomeni che caratterizzano il processo di ristandardizzazione dell'italiano, cui si è fatto cenno sopra, possono essere effettivamente interpretati in questo senso. In questa sede mi concentro principalmente sulla ridondanza sintagmatica. Un confronto tra la varietà standard-normativa e il cosiddetto italiano neostandard rivela numerosi casi in cui emerge una chiara divergenza nella ridondanza sintagmatica, con livelli di complessità crescente se si volge lo sguardo allo standard e livelli di complessità decrescente se invece l'obiettivo inquadra il neostandard. L'elenco di fenomeni neostandard che segue è indicativo e non esaustivo:

- (3) Mancato accordo tra verbo e soggetto posposto (Berruto 2012)
ce n'è di misteri in questa storia
- (4) Mancato accordo tra participio passato e oggetto sotto forma di pronome relativo (Sabatini 1985)
i libri che ho letto / le scarpe che luisa s'è comprata
- (5) Mancato accordo di genere e numero (Tempesta 1991)
tenendo presente le località
- (6) Coordinazione di preposizioni con ellissi di sintagma nominale uguale (Renzi 2012; Berruto 2017)
Traghetti da e per la Sardegna

In tutti questi casi, dunque, si osserva una riduzione di ridondanza sintagmatica rispetto alla varietà standard-normativa, che dunque si contraddistingue, rispetto a queste strutture, per un tasso maggiore di complessità.

Analoga conclusione si raggiungerebbe se si osservassero altri processi recenti associati al neostandard (ad esempio, la perdita di categorie morfologiche nel caso dell'avanzata di *loro* come soggetto in luogo

¹⁸ Per i quali rinvio a Berruto (1983).

di *essi /esse*, quindi senza flessione di genere; la riduzione di ridondanza paradigmatica, che di norma si realizza con l'affermazione di strutture analitiche in luogo di strutture sintetiche e che osserviamo, ad esempio, in forme del tipo *il ragazzo che gli ho prestato il motorino* rispetto a *il ragazzo cui ho prestato il motorino*; ecc.).

In sintesi, dunque, molti dei fenomeni che caratterizzano il processo di ristandardizzazione dell'italiano possono essere effettivamente interpretati come conferma del maggior livello di complessità della varietà standard-normativa e paiono confermare l'ipotesi avanzata sopra, secondo cui sia le varietà fortemente standardizzate all'interno di diasistemi complessi, sia le lingue 'piccole', soprattutto se isolate, esibiscono la tendenza a preservare, oltre a configurazioni tipologiche 'recessive', anche tratti e strutture di maggiore complessità¹⁹.

4. Conclusioni

Nei paragrafi precedenti si è mostrato come l'idea che il numero dei parlanti non debba essere considerato nella scelta delle lingue per un campione tipologico accantona, in modo un po' semplicistico, una serie di questioni che invece la tipologia dovrebbe affrontare, interagendo in modo più sistematico e strutturato con la sociolinguistica. Se è vero, infatti, che nessun tratto strutturale orienta o predetermina il successo di una lingua, è altrettanto vero che il numero di parlanti e le condizioni socio-geografiche in cui essi vivono possono incidere sul sistema della lingua, sia per quanto riguarda la configurazione globale del diasistema, sia per quanto concerne la maggiore o minore probabilità che alcuni *pattern* tipologici si affermino. In particolare, si è visto come nelle lingue 'piccole' e isolate tendano a stabilizzarsi tipi 'recessivi', cioè tipi statisticamente rari che, quindi, rischiano di sfuggire all'osservazione del tipologo. Per ridurre tale rischio, si è suggerito che queste lingue, nel campione, debbano essere sovrarappresentate. Abbiamo poi osservato come nei diasistemi complessi, tipici delle 'grandi' lingue del mondo nord-occidentale, possano convivere configurazioni tipologiche piuttosto distanti, che pongono un pro-

¹⁹ Trudgill (2011) ipotizza che la complessificazione possa manifestarsi soprattutto in lingue scarsamente apprese come L2, soprattutto da parlanti adulti. In effetti ciò caratterizza sia le lingue 'piccole', sia le varietà standardizzate nei diasistemi complessi.

blema metodologico significativo legato alla rappresentatività delle varietà rispetto all'intero diasistema. È evidente, infatti, come la scelta di assumere a riferimento una sola varietà per lingue molto complesse, come l'italiano, restituisca un'immagine parziale, sia proiettando sul diasistema tratti tipologici in realtà sociolinguisticamente marcati, sia 'nascondendo' tratti tipologici attestati in altre varietà. A questo problema si aggiunge una difficoltà ulteriore: in situazioni di questo tipo si tende, in modo quasi inerziale, a considerare come riferimento la varietà standard-normativa (se presente), in genere meglio descritta rispetto alle varietà non standard. Tuttavia, anche i processi di standardizzazione, soprattutto se di tipo *top-down*, hanno un grado significativo di 'innaturalità' e portano alla stabilizzazione di configurazioni tipologiche spesso simili a quelle delle lingue 'piccole'. In conclusione, quindi, se si trascura il numero di parlanti e le lingue vengono selezionate nella logica di 'uno vale uno' a prescindere, si corre il rischio di 'perdere' una porzione significativa di variazione, soprattutto nei tratti / tipi 'recessivi', sottovalutando il fatto che alcuni processi di fissazione di questi tratti avvengono, in modo tendenzialmente preferenziale, in lingue 'piccole' e isolate. Ma l'effetto prodotto dal combinato di una sovrarappresentazione delle lingue 'piccole' e dalla scelta di *default* della varietà standard / normativa, che è certo più e meglio descritta, ma spesso non prevalente nell'uso reale, in rappresentanza di diasistemi complessi rischia di ottenere l'effetto opposto, dando a tratti / tipi 'recessivi' un peso nettamente maggiore di quello che hanno in realtà!

Alla luce di queste considerazioni ci si può chiedere se, nei diasistemi complessi, le varietà non standard, soprattutto del parlato informale, con minor pressione normativa, cioè le varietà verosimilmente prevalenti nell'uso, possano essere considerate più rappresentative delle tendenze in atto nella lingua e quindi della sua configurazione tipologica. In questo caso, dunque, un campione tipologico dovrebbe comprendere un numero sovrarappresentato di lingue 'piccole', che non pongono problemi di variazione 'verticale', e le varietà non standard per i diasistemi complessi. Ciò che accomuna lingue 'piccole' e varietà non standard (soprattutto del parlato informale) dei diasistemi complessi è, molto semplicemente, il fatto che entrambe, pur in situazioni sociali radicalmente diverse, hanno parlanti: esse coprono, in effetti, la percentuale statisticamente preponderante dei membri di

una comunità e sono statisticamente prevalenti nell'uso quotidiano che essi fanno delle lingue. Dunque, restituiscono un'immagine delle lingue fedele alla realtà del loro impiego e alle loro 'condizioni di vita normali'²⁰.

Ringraziamenti

Ringrazio i due revisori per l'accurata rilettura del testo e per i generosi e preziosi suggerimenti, che hanno contribuito a migliorare un lavoro che resta, come usa dire, *in progress*. La responsabilità di ogni errore o imprecisione è esclusivamente mia.

Riferimenti bibliografici

- Bakker, Dik. 2012. Language Sampling. In Song, Jae Jung (a cura di), *The Oxford Handbook of Linguistic Typology*, 1-19. Oxford: Oxford University Press (prima ed. 2010).
- Ballarè, Silvia. 2020. L'italiano neo-standard oggi: stato dell'arte. *Italiano LinguaDue* 12(2). 472-492.
- Banfi, Emanuele. 2008. La famiglia delle lingue austronesiane. In: Banfi, Emanuele & Grandi, Nicola (a cura di), *Le lingue extraeuropee: Americhe, Australia e lingue di contatto*, 257-271. Roma: Carocci.
- Berretta, Monica. 2002. Correlazioni tipologiche fra tratti morfosintattici dell'italiano 'neo-standard'. In Dal Negro, Silvia & Mortara Garavelli, Bice (a cura di), *Monica Berretta. Temi e percorsi della linguistica. Scritti Scelti*, 379-410. Vercelli: Mercurio (ed. originale 1994).
- Berruto, Gaetano. 1983. L'italiano popolare e la semplificazione linguistica. *Vox romanica* 42. 39-79.
- Berruto, Gaetano. 2012. *Sociolinguistica dell'italiano contemporaneo*. Roma: Carocci (prima ed. 1987).
- Berruto, Gaetano. 2017. What is changing in Italian today? Phenomena of restandardization in syntax and morphology: an overview. In Cerruti, Massimo & Crocco, Claudia & Marzo, Stefania (a cura di), *Towards a*

²⁰ Una conseguenza metodologica di questa conclusione è quella di auspicare che la tipologia si apra, il più possibile, a studi *corpus-based* (di veda, ad esempio, Gerdes et al. 2021).

- new standard. Theoretical and Empirical Studies on the Restandardization of Italian*, 31-60. Berlin & New York: de Gruyter.
- Brugè, Laura & Brugger, Gerhard. 1996. On the Accusative 'A' in Spanish. *Probus* 8. 1-52.
- Cornilescu, Alexandra. 2000. Notes on the Interpretation of the Prepositional Accusative in Romanian. *Bucharest Working Papers in Linguistics* 2(1). 91-106.
- Cristofaro, Sonia. 2019. From synchronically-oriented typology to source-oriented typology. In: Gianollo, Chiara & Mauri, Caterina (a cura di), *CLUB Working Papers in Linguistics*, Volume 3. Bologna: Circolo Linguistico dell'Università di Bologna (<http://amsacta.unibo.it/6298/>).
- Dahl, Östen & Velupillai, Viveka. 2013. The Past Tense. In: Dryer, Matthew S. & Haspelmath, Martin (a cura di), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology (Available online at <http://wals.info/chapter/66>).
- Diessel, Holger. 2013. Distance Contrasts in Demonstratives. In: Dryer, Matthew S. & Haspelmath, Martin (a cura di), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology (Available online at <http://wals.info/chapter/41>).
- Dryer, Matthew S. 2013. Order of Subject, Object and Verb. In: Dryer, Matthew S. & Haspelmath, Martin (a cura di), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology (Available online at <http://wals.info/chapter/81>).
- Dryer, Matthew S. & Haspelmath, Martin (eds.). 2013. *The World Atlas of Language Structures Online*, Max Planck Institute for Evolutionary Anthropology, Leipzig (Available online at <http://wals.info>).
- Fiorentino, Giuliana. 2010. Accusativo preposizionale. In Simone, Raffaele (a cura di), *Enciclopedia dell'italiano*. Vol. 1, 17-19. Roma: Istituto dell'Enciclopedia Italiana - Treccani.
- Gerdes, Kim, Kahane, Sylvain & Chen, Xinying. 2021. Typometrics: From Implicational to Quantitative Universals in Word Order Typology. *Glossa: a journal of general linguistics* 6(1). 17.
- Gnerre, Maurizio. 2011. L'inafferrabile 'diversità' delle lingue. In Grandi, Nicola (a cura di), *Dialoghi sulle lingue e sul linguaggio*, 115-133. Bologna: Pàtron.

- Grandi, Nicola. 2013. Isolamento e insularità come parametri di analisi linguistica. *Evolution Day 2013 – Isole: laboratorio dell'evoluzione*. Milano: Museo di Storia Naturale (Comunicazione non pubblicata).
- Grandi, Nicola. 2020. La diversità inevitabile. La variazione linguistica tra tipologia e sociolinguistica. *Italiano LinguaDue* 12(1). 416-429.
- Himmelman, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: typological characteristics. In Adelaar, Alexander & Himmelman, Nikolaus P. (a cura di), *The Austronesian Languages of Asia and Madagascar*, 110-181. London & New York: Routledge.
- Levshina, Natalia. Token-based typology and word order entropy: A study based on Universal Dependencies. *Linguistic Typology* 23(3). 533-572.
- Nettle, Daniel. 1999. *Linguistic Diversity*. Oxford: Oxford University Press.
- Nettle, Daniel & Romaine, Suzanne. 2001. *Voci del silenzio. Sulle tracce delle lingue in via di estinzione*. Roma: Carocci (ed. originale 2000).
- Polinsky, Maria & Potsdam, Eric. In stampa. Austronesian Syntax. In: Palmer, Bill (ed.), *Languages of Oceania*. Berlin & New York: de Gruyter.
- Renzi, Lorenzo. 2012. *Come cambia la lingua. L'italiano in movimento*. Bologna: Il Mulino.
- Sabatini, Francesco. 1985. L'italiano dell'uso medio: una realtà tra le varietà linguistiche italiane. In Holtus, Günter & Radtke, Edgar (a cura di), *Gesprochenes Italienisch in Geschichte und Gegenwart*, 154-184. Tübinga: Narr.
- Sinnemäki, Kaius & Di Garbo, Francesca. 2018. Language Structures May Adapt to the Sociolinguistic Environment, but It Matters What and How You Count: A Typological Study of Verbal and Nominal Complexity. *Frontiers in Psychology* 9(1141).
- Song, Jae Jung. 2018. *Linguistic Typology*. Oxford: Oxford University Press.
- Tempesta, Immacolata. 1991. La relazione finale del seminario: un'occasione di apprendimento. In Lavinio, Cristina & Sobrero, Alberto A. (a cura di), *La lingua degli studenti universitari*, 47-70. Firenze: La Nuova Italia.
- Trudgill, Peter. 2011. *Sociolinguistic Typology. Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.

PARTE TERZA

COMUNICAZIONI

MARCO ANGSTER

Isolamento, cambio linguistico, integrità del sistema nel caso di una parlata in decadenza

Nel presente contributo si prenderà in considerazione una varietà alemanna parlata nell'Italia nord-occidentale (il *titsch* di Gressoney) riassumendone in primo luogo lo sviluppo socio-storico recente, dunque delineando una serie di fenomeni di cambiamento linguistico che in essa occorrono. L'insieme di questi fenomeni mostra come il sistema linguistico di questa varietà in decadenza, pur cambiando, mantenga sostanzialmente costante il numero e la tipologia di distinzioni morfosintattiche nel paradigma verbale. Il contatto non può essere facilmente accantonato come irrilevante in un campionamento tipologico in quanto le varietà che, come il *titsch*, si trovano in uno stato di decadenza e di forte contatto linguistico con una o più varietà dominanti sono soggette a un tasso di cambiamento linguistico più elevato di altre comunità più vitali e la stabilità tipologica del sistema linguistico è minore. Ciò rende possibile la compresenza in sincronia di diasistemi diversi tipologicamente e marcati sociolinguisticamente.

Parole chiave: minoranze walser, cambiamento linguistico, *Tun-Periphrase*, pronomi clitici, tipologia sociolinguistica.

1. Introduzione

I punti di contatto tra tipologia e sociolinguistica come messo in evidenza da Grandi (2019) sono evidenti. Entrambe le discipline sono studi della variazione: l'una interlinguistica, l'altra intralinguistica. Inoltre sia la sociolinguistica, sia la tipologia studiano le co-occorrenze di tratti con scopo descrittivo, ma anche predittivo. Citando Grandi (2019: 259):

Nel caso del tipo linguistico si fissa un principio funzionale e si 'prevede' l'occorrenza dei tratti linguistici ad esso associati (ad esempio, dicendo che l'italiano è testa iniziale si può prevedere che tipo di sintassi abbia). Nella varietà di lingua, si fissa una costante sociale e si 'prevede' l'occorrenza dei tratti linguistici ad essa associati (ad

esempio fissando la costante ‘popolare’ in diastria, si può prevedere che struttura abbiano le produzioni linguistiche dei parlanti che appartengono a quello strato della società).

Le due discipline si differenziano tuttavia anche profondamente: il diverso ambito di studio ha chiare conseguenze metodologiche. La tipologia per ottenere predizioni affidabili deve fondare le proprie generalizzazioni su di un campione di lingue che rappresenti la variazione interlinguistica evitando il più possibile deviazioni; Bakker (2011) ne identifica diversi tipi:

- bibliografica (*bibliographic bias*): gli studi tipologici si basano per lo più sulle descrizioni già esistenti e per moltissime lingue la documentazione o non esiste o è molto limitata;
- genetica (*genetic bias*): in parte risultante dalla deviazione bibliografica o dalla disponibilità di parlanti nativi o esperti di una lingua, un campione spesso sovrarappresenta le famiglie meglio documentate sottorappresentando le lingue di altre aree di elevata diversificazione genetica (Australia, Nuova Guinea, Sud America);
- areale (*areal bias*): un campione che non tiene conto dei fenomeni di contatto legati alle aree/leghe linguistiche può sovrarappresentare fenomeni che sono invece specifici di un’area linguistica e assenti o quasi altrove;
- tipologica (*typological bias*): un campione non deve sovrarappresentare determinati tipi linguistici soprattutto se l’oggetto della ricerca per cui il campione è costruito è influenzato dal fatto che una lingua appartenga a tale tipo;
- culturale (*cultural bias*): tale deviazione si lega al concetto di relativismo linguistico e alla possibilità che cultura e lingua si possano influenzare vicendevolmente – l’idea che culture diverse producano lingue diverse, ma anche che lingue e strutture linguistiche di un certo tipo influenzino il modo di concettualizzare la realtà.

Considerando queste deviazioni si può facilmente osservare come siano rilevanti per ciascuna di esse fattori esterni al linguaggio. È stato osservato ad esempio come lo studio dello Standard Average European condotto all’interno del progetto EUROTYP abbia privilegiato le varietà standard ignorando dialetti o varietà substandard (van der Auwera 2011: 299, 301). Privilegiare le varietà standard può essere una conseguenza diretta della distorsione bibliografica, perché esse

sono senz'altro dotate di una letteratura più ampia e approfondita rispetto alle varietà che sono spesso poco o affatto descritte, oppure i lavori su di esse si concentrano su singoli fenomeni, mentre mancano descrizioni puntuali e organiche del loro sistema linguistico. Ciò crea, se vogliamo, una distorsione bibliografica anche nella documentazione delle varietà che prescinde l'ambito tipologico e include invece quello sociolinguistico.

La distorsione genealogica dipende in parte come già detto dalla distorsione bibliografica, ma anche dal fatto, dovuto anch'esso a fattori extralinguistici – demografici, politici e sociali – che le lingue più conosciute sono anche le lingue per cui si hanno più parlanti; le lingue che hanno uno status più forte; le lingue per cui di più si investe in termini di produzione di risorse linguistiche. Anche in questo caso saranno dunque privilegiate lingue (indo)europee o lingue extra-europee maggiori, molte delle quali probabilmente coincidono con le lingue dotate di maggior documentazione.

Nel caso della distorsione tipologica il peso di fattori extralinguistici è più arduo da determinare, soprattutto in termini di costruzione di un campione tipologico. Esistono tuttavia proposte per collegare specifiche tendenze tipologiche (in particolare la semplificazione fonologica, morfologica e morfofonologica) a specifici tipi di comunità linguistiche sulla base di fattori quali: dimensione della comunità, struttura della rete sociale, stabilità sociale, contatto con altre comunità, conoscenza condivisa (Trudgill 2011). Se la semplificazione non è di per sé un tratto tipologico, vari tratti tipologici sono influenzati dalla complessità degli inventari di forme, si pensi solo, senza pretesa di completezza, alle seguenti carte del WALSL, dove le lingue campionate sono classificate secondo il numero di elementi che costituiscono i loro inventari fonologici o morfologici, oppure a seconda del tipo di neutralizzazioni tra forme diverse contenute nei paradigmi flessivi: 1A Consonant Inventories, 2A Vowel Quality Inventories, 22A Inflectional Synthesis of the Verb, 28A Case Syncretism, 29A Syncretism in Verbal Person/Number Marking, 30A Number of Genders, 49A Number of Cases, 58B Number of Possessive Nouns. Inoltre il passaggio di una lingua da un tipo morfologico sintetico ad uno analitico ha ovvie conseguenze sulla sintassi, che va a sopperire alla riduzione di distinzioni sul piano flessivo o in generale morfologico.

Più chiara è la rilevanza di fattori esterni al linguaggio nel caso della distorsione areale. Il criterio incorpora il concetto di contatto linguistico che a sua volta dipende da contingenze di tipo storico, politico e sociale. Oltre a ciò, seguendo Campbell (2006: 18), un'area linguistica non è che l'accumulo di casi specifici di diffusione localizzata di tratti linguistici. In precedenza (Campbell 1985) egli aveva contrastato nell'ambito dello studio dei fenomeni areali due approcci: uno "circostanzialista" (*circumstantialist*) – che semplicemente cataloga le somiglianze senza tentare di spiegarne la diffusione – e uno storicista (*historicist*) – che invece distingue tra i fenomeni realmente dovuti a diffusione (cioè al contatto linguistico) e le somiglianze accidentali (Campbell 1985: 32). È evidente come tali casi di diffusione localizzata di tratti linguistici possano essere seguiti soltanto attraverso un approccio storicista e un'attenzione diretta piuttosto alle varietà non-standard e parlate che non a quelle standard e scritte.

Discutendo le possibili distorsioni culturali – di per sé un esempio ulteriore della rilevanza dei fattori extralinguistici in ambito tipologico –, Bakker (2011: 6) suggerisce che le dimensioni di una comunità di parlanti possano essere un ulteriore parametro da considerare osservando come le lingue piccole abbiano una maggiore probabilità di contenere tratti tipologici esotici. A questo proposito Grandi (2019: 262) nota appunto che "la gamma di varietà che costituiscono un diaistema dipende in stretta misura dalla complessità della società cui fa riferimento". Ciò significa che se si considera una società grande ci si aspettano delle tendenze centrifughe esterne alla lingua con il risultato di una più alta variazione interna. Al contrario società piccole saranno piuttosto caratterizzate da tendenze centripete e la variazione interna sarà di conseguenza più ridotta.

Tuttavia, che cosa ci si può aspettare da una società piccola, ma sottoposta ad intense dinamiche di contatto come sono le lingue in decadenza? Nel mondo di oggi in cui la variazione linguistica è in pericolo e nella gran parte dei casi le lingue piccole sono anche lingue minacciate, non si può ignorare la possibilità che anche lingue piccole possano essere caratterizzate da un'elevata variazione interna. Già Dorian (1994: 694) osserva come una società estremamente omogenea come quella dei pescatori gaelici di Embo nell'East Sutherland è caratterizzata da un'elevata variazione – sia *inter-speaker*, sia *intra-speaker* –, collegando ciò alla mancanza (o scarsa adeguatezza) di una

varietà di prestigio a cui i parlanti possano riferirsi come norma. Un altro aspetto osservato da Dorian (1973: 437) è che il tasso di variazione nel gaelico parlato nella comunità di Embo appare molto più elevato di quello parlato in altre comunità, come Golspie o Brora, che più rapidamente hanno perso il proprio carattere gaelico e attribuisce questo effetto al prolungato periodo di bilinguismo inglese-gaelico trascorso però in un tendenziale isolamento.

Anche Trudgill (2011: 2) argomenta con una lunga serie di esempi che i cambiamenti che occorrono all'interno di una società partecipano a determinare o almeno a influenzare il tasso di cambiamento linguistico. In particolare mostrano di avere un ruolo da un lato il livello di contatto o di isolamento, dall'altro la relativa stabilità o instabilità sociale delle comunità: le varietà più conservative tendono ad essere relativamente più isolate geograficamente e relativamente più stabili socialmente rispetto a quelle più innovative (Trudgill 2011: 13-14). Trudgill inoltre osserva come comunità caratterizzate da alcuni specifici tratti sociali – quali basso livello di contatto linguistico tra adulti, elevata stabilità sociale, piccole dimensioni di una rete sociale densa e ampio corredo di conoscenze condivise – siano più favorevoli allo sviluppo di complessità linguistica (Trudgill 2011: 146).

In questo contributo si intende riprendere il problema della variazione interna ad una varietà in decadenza parlata nell'ambito di una piccola e (relativamente) compatta comunità di parlanti. La varietà considerata è il dialetto alemanno parlato a Gressoney (Valle d'Aosta), una minoranza linguistica tedesca storica tutelata nell'ambito della Legge 482/99. La comunità di Gressoney (così come quella di Issime, posta ad alcuni chilometri di distanza nella stessa valle del Lys) è caratterizzata da una situazione di contatto e di cambiamento linguistico (si veda oltre § 2.3) molto peculiare, perché si tratta di una comunità di minoranza tedesca incuneata nel continuum romanzo tra varietà gallo-italiche e francoprovenzale in una regione, la Valle d'Aosta, in cui il francese è lingua ufficiale accanto all'italiano, lingua ufficiale dello Stato italiano.

Nonostante le peculiarità della situazione del *titsch* di Gressoney siano di grande interesse per lo studio del contatto linguistico e più in generale della variazione linguistica in termini sia dialettologici, sia sociolinguistici, l'interesse di questa varietà dal punto di vista tipologico è limitato, almeno dal punto di vista della probabilità che il *titsch*

venga incluso in un campione tipologico di tipo tradizionale. Infatti, sia dal punto di vista genetico, sia tipologico, sia areale l'inclusione della varietà germanica parlata a Gressoney in un campione tipologico significherebbe scalzare dal campione stesso altre varietà europee e indoeuropee meglio e più ampiamente descritte come il tedesco standard o l'inglese.

Considerare il *titsch* in relazione al campionamento tipologico ha il valore non tanto di proporre questa varietà come un possibile candidato in un campione, quanto di mostrare quali siano gli effetti del contatto e della decadenza linguistica sulle strutture di una lingua piccola. Per il *titsch* però possediamo una documentazione relativamente ampia – soprattutto se confrontata a quella di lingue poco o per nulla descritte di famiglie esotiche. Disponendo inoltre di abbondante documentazione anche delle varietà ad essa affini (siano le altre parlate walser, i dialetti svizzero-tedeschi o tedeschi e il tedesco standard/letterario nelle sue varie fasi storiche), il livello di comprensione dell'evoluzione nel tempo della varietà è anche maggiore. In questo senso il presente contributo intende mostrare come anche una varietà piccola possa costituire un diasistema complesso.

Per fare ciò si intende rispondere alle seguenti domande:

- Il cambiamento linguistico incide sull'integrità del sistema (cfr. oltre § 3) linguistico del *titsch*?
- Quanto sono stabili i mutamenti nel sistema linguistico del *titsch*?
- Lungo quali dimensioni della variazione sociolinguistica si pongono le varianti generate dal cambiamento linguistico?

Il contributo è strutturato come segue: in § 2 si presenteranno le comunità walser italiane e verrà tratteggiato brevemente un loro quadro storico e sociolinguistico con speciale considerazione per il caso di Gressoney; in § 3 si discuterà del problema della conservazione dell'integrità del sistema linguistico del *titsch* di Gressoney con particolare attenzione all'evoluzione della morfologia verbale; in § 4 si sposterà l'attenzione su alcuni cambiamenti al confine tra morfologia e sintassi e si tenterà un'analisi della loro distribuzione nel diasistema del *titsch*; seguirà in § 5 una discussione sulla tensione tra mantenimento dell'integrità del sistema e cambiamento linguistico; nelle conclusioni (§ 6) si tornerà sul tema del rapporto tra tipologia e sociolin-

guistica suggerendo come le varietà in decadenza vadano considerate con attenzione nell'ambito del campionamento tipologico.

2. *I Walser in Italia*

2.1 Origine e colonizzazione

Le comunità walser italiane sono il residuo della colonizzazione di aree di alta quota delle valli alpine ad opera di coloni di lingua tedesco-alemana provenienti dal Vallese. Il nome stesso oggi usato per identificare le comunità e la sua origine sono rappresentativi dell'estensione della colonizzazione. Il termine *walser* è infatti l'esonimo in origine usato nelle aree più orientali raggiunte dai coloni per distinguere i gruppi vallesani (*Wallis* 'Vallese' < *wal(li)ser* 'vallesano') di parlata alemanna da quelli autoctoni di parlata bavarese.

La colonizzazione a sud delle Alpi è iniziata con la fondazione di Formazza, attestata a partire dal 1210 (Rizzi 2002) ed è proseguita nel corso del XIII secolo fino a comporre la costellazione di colonie solo in una certa misura territorialmente contigue che, come un arcipelago, emergono nel dominio romanzo – e in quello delle varietà alto alemanne, retoromanze o bavaresi a nord delle Alpi. La natura geografica della colonizzazione, che interessava in genere le parti di quota più elevata delle valli alpine, contribuì nei secoli a mantenere isolata ciascuna comunità di parlata (altissimo-)alemana sia dalle comunità di lingua romanza circostanti, sia dalle altre comunità linguisticamente e culturalmente affini. Al tempo stesso è noto come gran parte delle comunità walser a sud delle Alpi praticassero a partire almeno dal XVI secolo l'emigrazione stagionale o permanente verso l'area tedescofona a nord delle Alpi (Gressoney, Formazza, Alagna), ma anche verso le aree di lingua romanza sia a nord-ovest, in Francia (Issime), sia a sud delle Alpi, in Italia (Rimella). In tempi più recenti le comunità walser italiane hanno sperimentato, come molte località di montagna, processi di spopolamento, soprattutto nei casi in cui le località non abbiano sviluppato un solido turismo invernale connesso alle attività sciistiche.

2.2 Demografia

In Tabella 1 si riporta una stima del numero dei parlanti delle comunità walser italiane. I dati di popolazione si devono al censimento del 2001, mentre il numero di parlanti è stimato a partire dal rapporto CELE del 2002, che restituisce la percentuale del numero di abitanti che hanno risposto positivamente alla domanda “Conosce il Walser?”. Data la natura vaga della domanda, si intende il numero di parlanti senza dubbio come sovrastimato.

Tabella 1 - *Stima del numero dei parlanti delle comunità walser italiane. Gli ultimi parlanti attivi di Salecchio risiedevano fuori dal villaggio, oggi abbandonato*

	ISTAT (2001)	CELE (2002) %	“parlanti (stima)”
Gressoney-Saint-Jean	789	80.0	834
Gressoney-La-Trinité	297	68.3	
Macugnaga	651	56.7	369
Issime	403	79.8	322
Formazza	448	70.3	315
Rimella	142	90.8	129
Alagna	457	23.9	109
Salecchio	abbandonato	-	2
TOTALE	3187		2080

I dati demografici riportati in tabella 1 mostrano tutta la gravità della situazione delle parlate walser italiane. Il numero totale di abitanti in tutte le comunità di Piemonte e Valle d’Aosta non raggiungeva nel 2001 le 3200 unità, mentre, pur applicando le ottimistiche stime sul numero di abitanti che “conoscevano” nel 2002 la locale varietà alemanna si superano di poco i 2000 parlanti. Passati ormai 20 anni da quelle stime e sapendo che tendenzialmente in queste comunità la maggior parte dei parlanti attivi è anziana e la trasmissione intergenerazionale è scarsa, ci possiamo aspettare che oggi il numero di parlanti si sia ulteriormente ridotto.

Le comunità walser italiane costituiscono un esempio di lingue piccole e le singole comunità – che, va ricordato, di norma giacciono ciascuna nella parte più elevata di valli diverse – hanno costituito fino a tempi recenti delle società piccole linguisticamente e culturalmente disomogenee rispetto al territorio circostante.

2.3 Isolamento: fasi storiche

Vista la loro posizione geografica e la loro disomogeneità linguistica e culturale, le comunità walser sono state definite isole linguistiche (*Sprachinseln*) in area romanza (Zürrer 1982: 51)¹.

Considerando l'isolamento di queste comunità, soprattutto in termini diacronici, si pone il problema di definire da cosa precisamente e in quale periodo temporale esse siano isolate. Si può osservare infatti da un lato l'isolamento dall'area tedescofona, con conseguente persistenza di tratti arcaizzanti risalenti persino all'alto tedesco antico (Eufe e Mader 2018). Dall'altro lato si osserva anche quell'isolamento dall'area romanza che ha nel tempo contribuito al mantenimento per quasi otto secoli della peculiarità linguistica e culturale delle comunità stesse.

Gli aspetti dell'isolamento, nelle due forme menzionate sopra, evolvono però nel tempo e mutano al mutare delle dinamiche socio-economiche, politiche e culturali. Si possono in questo senso riconoscere alcune fasi storiche che caratterizzano le comunità walser italiane. Il quadro che presenteremo nel seguito di questo paragrafo va considerato come una generalizzazione e dunque una semplificazione delle vicende storiche delle comunità walser: data la loro differente dislocazione geografica e, di conseguenza, la loro appartenenza politico-amministrativa, esse presentano infatti differenze anche sostanziali dal punto di vista di isolamento e contatto. Inoltre, la ricostruzione privilegia il caso di comunità, quali Gressoney o Formazza, che fino a tempi abbastanza recenti hanno mantenuto contatti con le aree di lingua tedesca a Nord delle Alpi.

¹ In questa sede non si considererà come, almeno in termini di contiguità amministrativa tra i territori comunali occupati dalle comunità, alcune comunità walser italiane siano in effetti delle penisole linguistiche (si veda in proposito la classificazione proposta in Angster e Dal Negro 2017: 11). Inoltre, sempre Zürrer (1982) significativamente indica l'inizio della Prima guerra Mondiale come il momento a partire dal quale Gressoney “non è più inclusa nella periferia tedesca, ma, come isola linguistica, è separata e autoreferenziale” (Zürrer (1982: 51: *Seit die Handelsbeziehungen zur Schweiz und zu Deutschland mit dem Beginn des Ersten Weltkriegs abstarben, ist Gressoney nicht mehr in die deutsche Randzone einbezogen, sondern als Sprachinsel abgeschnitten und auf sich selbst verwiesen; [...]*). [Qui come in altri casi seguenti, se non esplicitato diversamente, la traduzione di estratti di opere in lingua straniera è da attribuirsi all'autore del contributo].

Al netto di tali premesse, possiamo suddividere la storia delle comunità walser italiane in quattro fasi storiche (si veda per maggiori dettagli Angster, Gaeta in stampa: 85-88 e i riferimenti in Angster 2012: 162-164):

- 1) fase della fondazione e colonizzazione (secoli XIII-XIV);
- 2) fase delle migrazioni stagionali (verso Svizzera, Germania o Francia; dal tardo XV al tardo XIX secolo);
- 3) fase del Fascismo (essenzialmente tra fine Prima e fine Seconda Guerra Mondiale);
- 4) dal Secondo Dopoguerra a oggi.

In ciascuna delle fasi le dinamiche di isolamento e contatto cambiano (e, come detto, possono differire anche radicalmente tra le diverse comunità). Si può tuttavia considerare che, soprattutto nella seconda fase e per quelle comunità per le quali la migrazione era diretta soprattutto verso Svizzera (tedesca) e Germania meridionale, l'isolamento dall'area romanza – inteso come scarsa permeabilità alla penetrazione di tratti culturali esterni – sia stato più forte in quest'epoca in cui, per contro, il legame economico e culturale con l'area tedescofona si è mantenuto a lungo vitale². L'isolamento dall'area romanza non esclude l'ingresso nella comunità di individui provenienti dal suo esterno, ma il numero di ingressi e la capacità di assimilazione erano tali da non modificare significativamente il profilo sociale e dunque linguistico e culturale della comunità³.

² A testimonianza di questa vicinanza culturale con l'area tedescofona si può menzionare almeno Schott (1842) che nota come «a Gressoney, di cui scuola e chiesa da lunghi anni sono tedesche e i cui contatti sono in gran parte con la Germania, per ogni uomo il tedesco è del tutto familiare» (Schott 1842: 135: [...] *in Gressoney, dessen schule und kirche seit langen jahren deutsch sind, dessen verbindungen gröstentheils nach Deutschland geben, fast jedem mann das Hoch-deutsche vollkommen geläufig ist.*). Già Schott testimonia invece una situazione ben diversa a Issime, dove il dialetto gli è incomprendibile «perché la scuola è francese e perché il commercio degli issimesi non è indirizzato verso la Germania e la Svizzera. Ho dovuto quindi parlare con le persone in francese.» (Schott 1842: 13: [...] *weil die schule französisch und weil der verkehr der Issimer nicht nach Deutschland und der Schweiz gerichtet ist. Ich musste also mit dem leuten französisch reden.*).

³ Si vedano in questo senso i numerosi casi di famiglie che assimilano individui originari dell'area romanza a Gressoney (Zürner 1982: 47) o che cambiano la propria appartenenza linguistica nella zona di transizione costituita dai territori di Gaby e Issime (Musso 2017: 13-14).

Il quadro si modifica gradualmente tra l'Unità d'Italia e la fine della Prima Guerra Mondiale, quando, con le mutate condizioni geopolitiche e il successivo avvento del Fascismo in Italia, le tradizionali dinamiche migratorie si interrompono e l'atteggiamento verso le minoranze alloglotte si irrigidisce per la politica di italianizzazione promossa dal regime.

Con la fine del Fascismo e della Seconda Guerra Mondiale, se da un lato vengono meno le limitazioni imposte alle comunità di minoranza dalla politica linguistica fascista, dall'altro i cambiamenti sociali, tecnologici e culturali in seno alla società vanno ad influire capillarmente sulla vita quotidiana delle persone anche nelle aree più remote. Il ruolo dell'italiano si rafforza, come le spinte centrifughe dello spopolamento e dell'eterogamia.

Solo negli ultimi decenni le varietà walser hanno goduto di un aumento di prestigio e considerazione grazie al lavoro delle associazioni culturali che sono state fondate in seno alle comunità con lo scopo di preservarle. In questo contesto il termine *walser* è stato usato per la prima volta come denominazione che riunisce tutte le comunità di tale origine sotto un'unica identità.

Riassumendo si può notare come il cambiamento più vertiginoso si ha nel corso delle ultime due fasi, corrispondenti circa agli ultimi 100 anni. Se concentriamo l'attenzione su Gressoney, i cambiamenti di quest'ultimo secolo coinvolgono i seguenti aspetti:

- l'interruzione dei secolari movimenti migratori stagionali verso Svizzera e Germania;
- l'obbligo dello studio dell'italiano (imposti lentamente a partire dall'introduzione nel 1859 dell'obbligo scolastico; efficace probabilmente solo dal 1923 quando un Regio decreto stabilisce che l'italiano sia la lingua esclusiva di insegnamento; affiancato dal francese in Valle d'Aosta a partire dal 1945);
- il dilagare dell'italiano (accanto al piemontese e in misura minore al francoprovenzale) come lingua parlata e di socializzazione primaria per la mancata assimilazione dei nuovi arrivati;
- abbandono del tedesco letterario come lingua di alfabetizzazione e "riferimento normativo";
- transizione verso lo scritto del *titsch*⁴.

⁴ Sull'evoluzione del repertorio linguistico a Gressoney con particolare riferimento alle lingue studiate a scuola, si veda Angster (2014).

I cambiamenti elencati qui sopra mostrano come il parametro dell'isolamento a Gressoney si ribalti. Se fino a inizio XX secolo si può sostenere che vigesse un tendenziale isolamento culturale rispetto all'area romanza, all'inizio del XX secolo nel giro di pochi anni si perviene ad un ben più rigido isolamento culturale rispetto all'area tedescofona.

In termini di repertorio linguistico va notato che Gressoney nel XIX secolo mostra i tratti di una diglossia tra dialetto locale (*titsch*) come varietà bassa (*L variety*) e tedesco letterario (localmente denominato *bòchtitsch* 'alto tedesco' o *guettitsch* 'buon tedesco') come varietà alta (*H variety*) (Zürrer 2009: 86). La situazione evolve fino a vedere l'ingresso durante il '900 prima del piemontese come varietà bassa accanto al *titsch* e poi dell'italiano come varietà alta accanto al tedesco letterario (Zürrer 2009: 100)⁵. In seguito il repertorio si arricchisce ancora con il francese come varietà alta, mentre tra le varietà basse si impone una dilalia tra varietà locale germanica (accanto al piemontese) e italiano⁶.

Le conseguenze di questi sviluppi sociali e culturali sono fatali per la sopravvivenza delle varietà. Prima però di determinarne la scomparsa, esse ne accelerano il cambiamento: il venir meno del riferimento culturale tedesco "libera" infatti del tutto il *titsch* dalla pressione normativa della lingua tedesca letteraria e determina quindi per esso quella condizione di equilibrio tra il cambiamento linguistico – indotto dal contatto (se si considerano le varietà romanze), ma anche dalla sua assenza (se si pensa al tedesco letterario) – e il completo abbandono della varietà ancestrale a favore dell'italiano.

⁵ Va notato il contrasto con la vicina comunità di Issime, dove già nel XIX secolo si attesta l'uso del francoprovenzale come varietà bassa e del francese come lingua tetto per il francoprovenzale, mentre la varietà germanica locale, il *töitschu*, per quanto ci è noto appare priva di copertura (*Dachlos*) lungo tutta la propria storia recente (si veda anche Zürrer 2009: 87, 100).

⁶ Definire in che momento le varietà romanze si inseriscano nel repertorio della comunità di Gressoney come varietà bassa è difficile da stabilire, si può tuttavia supporre che ciò corrisponda al momento in cui in una certa fascia d'età il numero di parlanti competenti diviene minoranza. Dai dati di Giacalone Ramat (1979) e Squinabol (2008) discussi in Angster (2014: 114-116) questo momento sembrerebbe occorrere dopo la metà degli anni '60 del '900.

Ci possiamo chiedere a questo punto quale posizione occupi il *titsch* di Gressoney nella proposta di Trudgill (2011: 147) di combinazione – qui riportata in Tabella 6 – dei tre fattori sociali di *dimensione* (grande/piccola), *rete (sociale)* (stretta, ampia) e *contatto (linguistico)* (basso/elevato).

Tabella 2 - *Tipologia delle comunità linguistiche basata su fattori sociali*
(Trudgill 2011: 147)

	1	2	3	4	5	6
<i>size</i>	small	small	small	small	large	large
<i>network</i>	tight	tight	loose	loose	loose	loose
<i>contact</i>	low	high	low	high	low	high

Nella sua tipologia di sei categorie ottenute incrociando i valori di questi tre tratti possiamo posizionare la comunità di Gressoney per quanto riguarda la sua situazione fino alla Prima Guerra Mondiale nella *categoria 1 – comunità piccola, dalla rete sociale stretta, sottoposta ad un contatto basso*. Il contatto aumenta (o comunque cambia direzione e cresce in intensità) nel periodo tra le due guerre e fino al primo dopoguerra facendo passare la comunità nella *categoria 2 – comunità piccola, dalla rete sociale stretta, sottoposta ad un contatto elevato*. Ancora nel primo periodo successivo alla Seconda Guerra Mondiale si può considerare la comunità come dotata di una stretta rete sociale, ma tale tratto cambia e probabilmente dopo la seconda metà degli anni '60 del secolo scorso (si veda nota 6 sopra) la comunità finisce nella *categoria 4 – comunità piccola, dalla rete sociale larga, sottoposta ad un contatto elevato*.

2.3 I dati

Nonostante il contesto di decadenza linguistica tratteggiato sopra, non si considererà qui il caso dei semi-parlanti. Si considererà invece il *titsch* nelle produzioni di parlanti con (almeno ipoteticamente) piena competenza.

I dati che verranno discussi in quanto segue provengono – oltre che dalla letteratura secondaria – da diverse fonti pubblicate nell'ambito di varie iniziative di documentazione linguistica che hanno interessato le minoranze walser a Sud delle Alpi o in generale le minoranze tedesche storiche in Italia. Nello specifico sono tratti dai corpora

scritti raccolti ed elaborati nel contesto del progetto *ArchiWals* (abbreviato in seguito negli esempi come CAW; Angster et al. 2017); dalla raccolta di frasi presenti nel *Tesoro linguistico delle isole germaniche in Italia* tradotte dall'italiano verso tutte le varietà tedesche di minoranza rappresentate nel *Comitato unitario delle isole tedesche storiche in Italia* (in seguito TESORO; Geyer et al. 2014); dagli etnotesti che corredano le carte geolinguistiche raccolte nel *Piccolo Atlante Linguistico dei Walser Meridionali* (in seguito PALWAM; Antonietti et al. 2015).

3. Integrità

Dopo la panoramica sulle comunità walser a Sud delle Alpi esposta negli scorsi paragrafi, rivolgiamo ora l'attenzione ad alcuni fenomeni che caratterizzano le varietà considerate, con speciale attenzione alla varietà di Gressoney. Considereremo dunque se e in che misura i cambiamenti sociali, politici e culturali menzionati in precedenza siano correlati a un cambiamento nelle strutture delle varietà walser e valuteremo innanzitutto il problema dell'integrità del sistema linguistico.

Per integrità intendiamo qui la persistenza nel sistema di tratti linguistici, come il mantenimento e la mancata riduzione di distinzioni grammaticali, presenti in una fase antecedente ai mutamenti extralinguistici (cioè fattori esterni di tipo sociale, demografico ecc.). La nostra tesi è che, nonostante i mutamenti extralinguistici e il cambio linguistico osservabile nell'ultimo secolo, il sistema linguistico del *titsch* mantenga la propria integrità. Come detto in precedenza, non si considera in questo senso il caso dei semi-parlanti, per i quali ci si può aspettare un quadro differente (si veda in proposito Dal Negro 2004 per il caso di Formazza). Per dimostrare la nostra tesi considereremo il caso della flessione verbale, innanzitutto dal punto di vista dell'evoluzione dei paradigmi delle classi di verbi.

3.1 Classi di verbi

Nel corso del XX secolo è possibile osservare cambiamenti evidenti nella flessione verbale del *titsch* attestati dalle descrizioni di Bohnenberger (1913), Zürrer (1982) e dai paradigmi inclusi nei vocabolari pubblicati dal *Centro Culturale Walser / Walser Kulturzentrum* (WKZ 1988).

Vediamo in Tabella 2 le forme dell'infinito, del presente indicativo e del participio perfetto del *Kurzverb* 'vedere', del modale preterito-presente 'potere' e del verbo non-ridotto 'viaggiare'⁷.

Tabella 3 - *Classi di verbi in titsch (1)*

	Inf.	Ind. pres.			1.-3. Pl.	Part. Perf.
		1. Sg.	2. Sg.	3. Sg.		
a. Kurzverb (verbo contratto): 'vedere'	<i>gsē</i>	<i>gsēn</i>	<i>gsēšt</i>	<i>gsēt</i>	<i>gsēχen</i>	<i>gsēt</i>
'lasciare'	<i>loa</i>	<i>lān</i>	<i>lāšt</i>	<i>lāt</i>	<i>lein</i>	<i>gloa/gloat</i>
b. modale (preterito-presente): 'potere'	<i>χonnu</i>	<i>χan</i>	<i>χanšt</i>	<i>χant</i>	<i>χannen</i>	<i>χonnu</i>
c. verbo non-ridotto: 'viaggiare'	<i>foare</i>	<i>foaren</i>	<i>foaršt</i>	<i>foart</i>	<i>foaren</i>	<i>gfoaret</i>

Angster, Gaeta (2018), analizzando l'evoluzione del gruppo dei *Kurzverben* e dei modali lungo il '900, mostrano che tra le attestazioni più antiche e le più recenti la flessione di questi verbi resta in sostanza stabile. Si attestano tuttavia delle innovazioni che, benché marginali, portano verso una stabilizzazione delle condizioni per cui un verbo appartiene al gruppo dei *Kurzverben*:

- la presenza di forme monosillabiche nel singolare del presente indicativo, dell'infinito e del participio perfetto;
- la differenziazione formale tra forme singolari e plurali al presente indicativo (legata al cambio della vocale radicale o alla riemersione delle consonanti finali della radice).

Tali condizioni non sussistono invece per il gruppo dei modali (o, più correttamente, dei preterito-presenti) che appare invece caratterizzato dal mantenimento della consonante finale della radice e – anche se appare questo un tratto in recessione – dal cambio della vocale radicale tra singolare e plurale (es. *wil* 'volere.IND.PRS.1PL' ~ *wellen* 'volere.IND.PRS.1PL'; *muß* 'dovere.IND.PRS.1PL' ~ *mussen/missen* 'dovere.IND.PRS.1PL'). Il gruppo, più ricco di lessemi, dei *Kurzverben*

⁷ I *Kurzverben* e i preterito-presenti sono gruppi di verbi che si definiscono per lo più sulla base di ragioni formali e non semantiche. In particolare i *Kurzverben* non provengono storicamente da stesse classi flessive di fasi più antiche del tedesco. Per esempio *loa*, pur facendo parte del novero dei *Kurzverben*, ricopre varie funzioni grammaticali, tra cui il causativo (permissivo). Al contrario, il verbo *wissò* 'sapere' non ha alcuna funzione grammaticale, ma partecipa delle stesse caratteristiche formali degli altri preterito-presenti, che però dal punto di vista funzionale sono tutti verbi modali.

sembra però attrarre verso le proprie caratteristiche formali almeno un modale, *mögu* ‘aver voglia, volere’ (ted. *mögen*). Tale verbo appare assimilarsi alle caratteristiche del gruppo dei *Kurzverben* tanto che la sua flessione viene a ricalcare nell’indicativo presente singolare quella di *loa* ‘lasciare’ (ted. *lassen*).

Tabella 4 - *Classi di verbi in titsch* (2)

	Inf.	Ind. pres.			1.-3. Pl.	Part. Perf.
		1. Sg.	2. Sg.	3. Sg.		
a. Kurzverb (verbo contratto): ‘lasciare’	<i>loa</i>	<i>lān</i>	<i>lāst</i>	<i>lāt</i>	<i>lein</i>	<i>gloa/gloat</i>
b. modale (> contratto): ‘aver voglia, volere’	<i>mögu</i>	<i>mān</i>	<i>māst</i>	<i>māt</i>	<i>mein</i>	<i>mögu</i>
c. modale (preterito-presente): ‘potere’	<i>χonnu</i>	<i>χan</i>	<i>χanšt</i>	<i>χant</i>	<i>χannen</i>	<i>χonnu</i>

Il caso delle classi flessive, in particolare di *Kurzverben* e modali, mostra come gli effetti del cambio linguistico non portino necessariamente ad un livellamento analogico verso le classi verbali non ridotte, morfologicamente più regolari. Il caso di *mögu*, attratto alle caratteristiche dei *Kurzverben*, mostra come questa classe dalle caratteristiche idiosincratiche possa tuttavia fungere da modello analogico.

3.2 Sviluppo di forme verbali analitiche

La discussione in § 3.1 considera esclusivamente le forme sintetiche dei paradigmi verbali. Inoltre si basa su dati raccolti al più tardi negli anni ’70-’80 del secolo scorso. Già nelle fonti più recenti considerate (ad es. Zürner 1982) si fa però riferimento alla marcata tendenza alla sostituzione delle forme verbali sintetiche con forme analitiche costruite con il verbo *tue* ‘fare’ (ted. *tun*) che regge l’infinito del verbo. Vediamo qui in (1) un esempio di ciò al presente.

- (1) *D’* *Bur-e* *tien* *jetza* *andersch*
 DEF.NOM.PL contadino-PL fare.3PL ora diversamente
buoro.
 coltivare.INF
 ‘I contadini ora coltivano diversamente’ [CAW]

Va chiarito a questo proposito che, come nella gran parte delle varietà di tedesco superiore, l’inventario di forme verbali sintetiche nel *titsch* non include il preterito: forme di preterito infatti non sono attesta-

te neppure nelle descrizioni più antiche disponibili⁸. L'espressione di azioni passate si ottiene dunque esclusivamente tramite il perfetto analitico costruito con gli ausiliari *si* 'essere' o *hā* 'avere' (ted. *sein* e *haben*, rispettivamente) a reggere il participio perfetto del verbo.

Nel *titsch* di Gressoney la perifrasi con *tue* 'fare' (*Tun-Periphrase*) e la perdita del preterito sintetico (*Präteritumschwund*) collaborano ad una radicale innovazione del paradigma verbale in senso analitico. I due fenomeni sono tuttavia ampiamente attestati anche in altre varietà di tedesco (e in generale germaniche).

La *Tun-Periphrase* "serpeggia" come tratto non-standard in tedesco da secoli⁹. Langer (2000) ne mostra l'elevata frequenza in alto tedesco protomoderno (1300-1650) e la diffusione in tale periodo (così come oggi) in tutte le aree dialettali tedesche. La sua assenza nello standard – se non in casi marginali, ad esempio come proforma o per topicalizzare il predicato – si deve ad un processo di stigmatizzazione del suo uso che lo ha portato nel XVIII secolo ad essere considerato un tratto del parlato degli strati bassi della popolazione (Langer 2000: 314).

Al contrario, il *Präteritumschwund* caratterizza più specificamente le varietà di tedesco superiore, dove si può osservare la crescita dell'uso delle forme di perfetto a partire dal periodo medio alto tedesco, con un picco tra XVI e XVII secolo (dunque proprio al termine del periodo protomoderno). Successivamente, per il rafforzamento dell'influsso della lingua scritta, il preterito ritorna in auge in area tedesca meridionale, almeno nella lingua scritta (Fischer 2018: 159). È significativo menzionare il fatto che, a fronte di una generale riduzione delle forme di preterito attestate e di una tendenza del preterito a conservarsi tanto meglio quanto più a nord una varietà dialettale è parlata, si

⁸ Già Bohnenberger (1913) testimonia della completa scomparsa dell'indicativo preterito in tutta l'area delle parlate vallesane: «La coniugazione sembra aver perso l'indicativo del preterito in tutta l'area, così che essa possiede ormai soltanto un indicativo nel verbo non composto.» [Die Konjugation scheint im ganzen Gebiete den Indikativ des Präteritums verloren zu haben, so daß sie im nicht zusammengesetzten Verbum nur noch einen Indikativ besitzt.] (Bohnenberger 1913: 224).

⁹ Si veda però Hill (2010) per una discussione del preterito debole germanico a dimostrazione dell'uso di 'fare' come ausiliare addirittura nella fase protogermanica: «The Germanic weak preterite is a periphrastic formation which consists of a verbal noun and grammaticalized inflections of an old verbal stem belonging to the verbal root meaning "do"» (Hill 2010: 451).

può tuttavia evidenziare una resistenza del preterito “verbo-specifica”. Fischer (2018: 390) elenca in particolare i seguenti verbi: *sein, haben, wollen, sollen, können, müssen, dürfen, sagen, wissen, kommen, denken, geben, gehen, stehen, werden, nehmen, sitzen, tun*. L’autrice nota che sono “verbi forti e irregolari”, molti di essi sono “verbi copulari, modali o ausiliari sintatticamente funzionalizzati” e “verbi frequenti che appartengono al lessico di base e che in molti casi presentano una semantica verbale imperfettiva”¹⁰.

Ad una resistenza verbo-specifica del preterito va pure collegata la sopravvivenza, come evidenziato da Nübling (1997) specialmente nei dialetti svizzero tedeschi, del *Konjunktiv II*, cioè di quell’insieme di forme che serve all’espressione della potenzialità, dell’irrealtà e del desiderio (ottativo). Ciò che appare particolarmente significativo è che il *Konjunktiv II* (es. ted. *ich nähme* ‘prenderei’) dipende formalmente, cioè si forma sulla base del preterito (cfr. ted. *ich nahm* ‘presi’), del quale ritiene in alto tedesco antico il vocalismo – poi successivamente palatalizzato in alto tedesco medio per metaforesi (*Umlaut*) condizionato dalle desinenze contenenti /i/ specifiche del congiuntivo (ata. *nām-i* > mat. *neme*). La sua sopravvivenza è dunque contraria alle attese, vista la contemporanea totale scomparsa del preterito nelle varietà alemanne e tedesco superiori in generale (Nübling 1997: 107).

Così, tornando al *titsch* di Gressoney, se da un lato il preterito è del tutto assente dalle attestazioni di questa varietà, dall’altro almeno ancora nella descrizione in Zürcher (1982: 94-96) si attestano forme di *Konjunktiv II* sintetiche formate con la desinenza debole *-ti*. Va notato come le forme deboli caratterizzino in *titsch* anche ausiliari (Inf. *hā* ‘avere’ > Konj-II *hetti* ‘avessi’), modali (Inf. *χonnu* ‘potere’ > Konj-II *χanti* ‘potessi’), e *Kurzverben* (Inf. *gse* ‘vedere’ > Konj-II *gsexti* ‘vedessi’). Unico verbo a mantenere una forma forte, cioè *ueri* ‘fossi’, è *si* ‘essere’, che però accanto ad essa ne presenta anche una debole: *werti* ‘fossi’.

Il *titsch* non ignora la possibilità dell’utilizzo di forme analitiche di *Konjunktiv II*¹¹. Anche in questo caso, come si è visto sopra per il pre-

¹⁰ «Es sind in erster Linie irreguläre und starke Verben. Viele von ihnen sind als Kopula-, Modal-, oder Hilfsverben syntaktisch funktionalisiert. Es sind häufige Verben, die zum Grundwortschatz gehören und von denen viele eine imperfektive Verbsemantik haben» (Fischer 2018: 390).

¹¹ Sempre Bohnenberger (1913) menziona questa possibilità come sempre più frequente in tutte le parlate vallesane: «Fortemente ridotta è anche la formazione del

terito, l'espansione delle forme analitiche è verbo-specifica e occorre nei casi in cui un verbo ausiliare al *Konjunktiv II* regge l'infinito di un altro verbo per cui una forma sintetica non è disponibile. Proprio in questo contesto la *Tun-Periphrase*, con la forma debole di *Konjunktiv II teti* 'facessi', viene a sopperire all'obsolescenza delle forme sintetiche di altri verbi: *ix laxti* > *ix teti laxe* 'ridessi'. In questi casi la *Tun-Periphrase* costituisce oggi l'unica scelta possibile per la gran parte dei verbi di bassa frequenza, con la sola esclusione di ausiliari e modali (per quanto concerne i *Kurzverben*, si veda oltre).

Tuttavia la *Tun-Periphrase* può sostituire di fatto ogni forma sintetica del paradigma verbale con un corrispondente analitico, come si può osservare dagli esempi seguenti dove forme sintetiche di *tue* 'fare' sostituiscono di volta in volta le forme sintetiche di presente indicativo visto sopra in (1), imperativo in (2), congiuntivo presente (cioè *Konjunktiv I*) in (3) o congiuntivo preterito/condizionale (*Konjunktiv II*) in (4):

- (2) *Du vom* *Himmel, tue* *pschetze*
 2sg da.DEF.M.DAT.SG cielo fare.IMP.2SG proteggere.INF
enz *Land.*
 POSS.1PL.N.NOM terra
 'tu, dal cielo, proteggi la nostra terra' [CAW]
- (3) *Dass z'* *Chrésch-ként-le* *tiege*
 COMP DEF.N.NOM.SG Cristo-bambino-DIM fare.SUBJ.PRES.3SG
gscheng-e *brénge*
 regalo-PL portare-INF
éscht *an* *schen-e*
 essere.IND.PRES.3SG INDF.M.NOM.SG bello.M.NOM.SG
bruch...
 usanza(M)
 'Che Gesù bambino porti i doni è una bella usanza' [CAW]
- (4) *Hie allz* *uf-z-schrib-e,* *was* *Eigen*
 qui tutto.N.NOM.SG elencare-COMP-STEM-INF che.cosa Eugenio
hät *toat,* *weré*
 avere.IND.PRES.3SG fare.PART.PRT essere.SUBJ.PRT.3SG

congiuntivo del preterito, se la perifrasi con 'fare' è usata frequentemente» [Stark eingeschränkt ist auch die Bildung des Konjunktivs des Präteritums, sofern häufig Umschreibung mit „tun“ angewandt wird] (Bohnenberger 1913: 224).

schier ònmégléch òn mò tette sécher
 quasi impossibile e IMPR fare.SUBJ.PRT.3SG di.certo

mengs déng vergesse.
 molto.N.NOM.SG cosa(N) dimenticare.INF

‘Elencare qui tutto ciò che Eugenio [...] ha fatto sarebbe quasi impossibile e si dimenticherebbero di certo molte cose’ [CAW]

Si noti negli esempi (1)-(4) che *tue* non contribuisce semanticamente al significato complessivo delle frasi fungendo solo da *locus* della marcatura dei tratti TAM e di persona.

Ciò che è importante notare è come, nonostante il dilagare di forme analitiche per tutti i verbi con l’eccezione di ausiliari (‘essere’, ‘avere’ e ‘fare’), modali e (alcuni) altri *Kurzverben*, le distinzioni nel paradigma verbale si mantengono. *Konjunktiv I e II*, benché espressi in forma analitica continuano a far parte dell’inventario di distinzioni formali e di funzioni della flessione verbale del *titsch*.

Anche la *Tun-Periphrase* inoltre si applica secondo un’espansione verbo-specifica: gli ausiliari *si* e *hä* mantengono attive tutte le forme sintetiche dal presente al *Konjunktiv II*. Soltanto *tue* ‘fare’ curiosamente consente la scelta tra forma sintetica e forma analitica raddoppiata: *iχ teti / iχ teti tue*. Analogamente gli altri *Kurzverben* non-ausiliari sono compatibili con la *Tun-Periphrase*: appare anzi l’unica scelta nel caso del *Konjunktiv II* (le cui forme sono però attestate in Zürrer 1982).

3.3 Distribuzione delle forme analitiche e resistenza delle forme sintetiche

Possiamo vedere quanto possa essere pervasiva la *Tun-Periphrase* nella trascrizione di parlato dal seguente etnotesto – pubblicato nel *PALWaM*, Antonietti *et al.* 2015 – dove le forme di *tue* che reggono un verbo all’infinito sono sottolineate nel brano (tradotto in nota 10).

(5)

VC: eh, dunque, *du wefëll chie häscht?*

L: *hännéró só fënföntzwentzg en allem aber z mälche elwé*

VC: *elwé jetza*

L: *jetza, ja, ja òn em sömmer es bétzié mé wóróm gämber óf z alpó de*

VC: certo

L: *tieberó gé es pare z zueft òn òn sirró es bétzie mé*

- VC: *ón em sómmer tuescht óu machó de chésch?*
 L: *ja*
 VC: *anschtatt jetza es bétz wenégor*
 L: *ja, ja ma jetza schier néks fórom sinntsch d chalbiene ón*
 VC: *ón tién dšchi trénge d mélch*
 L: *de tién dšchi trénge d mélch bés wenn sinn gmaschté ón de tieber,
 de tuené de afoa z chéschó*
 VC: *ón wewéll moal z tagsch tuescht mälché?*
 L: *zwei moal z tagsch de morgé ón em oabe*
 VC: *ón wewéll mélch? Tientsch machó?*
 L: *ma óngéfer d chue wie*
 VC: *certo fón*
 L: *fón dri litter só bés zwelfé, dritzené. [PALWAM: 43]¹²*

Il brano di intervista presenta 18 turni e in essi si riconoscono 14 forme verbali finite di cui 8 presentano la *Tun-Periphrase* (si conta qui anche il caso di *tieber* (*afoa*) ‘cominciamo’, poi ripianificato in *tuené afoa* ‘comincio’). I verbi retti da *tue* sono i seguenti: *machó* ‘fare’ (2 occorrenze), *trénge* ‘bere’ (2), *mälche* ‘mungere’ (1); inoltre due *Kurzverben* hanno forme analitiche rette da *tue*: *gé* ‘prendere’, *afoa* ‘incominciare’ (derivato di *foa* ‘catturare’, cfr. ted. *fangen*) ciascuno con una occorrenza. Sono soltanto sei le forme verbali sintetiche attestate nel brano: cinque riguardano gli ausiliari *si* ‘essere’ (3 occorrenze) e *há* ‘avere’ (2), e una il *Kurzverb* *goa* ‘andare’¹³.

Il quadro che emerge dal pur breve brano in (5) concorda con le conclusioni proposte in Angster (2011) – e riassunte qui sotto in Tabella 5 – sulla base di un test di accettabilità somministrato a quattro parlanti di età diverse¹⁴.

¹² «VC: eh, dunque, tu quante mucche hai? / L: ne ho all’incirca venticinque in tutto, ma da mungere undici / VC: undici adesso / L: adesso, sì, sì e in estate un pochino di più perché andiamo in alpeggio / VC: certo / L: ne prendiamo alcune in affitto e ce ne sono un pochino di più / VC: e d’estate fai anche il formaggio? / L: sì / VC: invece adesso un po’ meno? / L: sì, sì ma ma adesso quasi niente perché ci sono i vitellini e / VC: e loro bevono il latte / L: loro bevono il latte fino a quando sono cresciuti e poi noi, allora inizio a fare il formaggio / VC: e quante volte al giorno mungi? / L: due volte al giorno alla mattina e alla sera / VC: e quanto latte? fanno? / L: ma a seconda della vacca / VC: certo da / L: da tre litri così fino a dodici tredici.» (Antonietti et al. 2015: 43, con modifiche).

¹³ Si noti che *goa* ‘andare’ funge da ausiliare nella costruzione passiva su cui si veda Gaeta (2018).

¹⁴ I dati discussi in Angster (2011) derivano dal lavoro sul campo alla base di Angster (2004/2005) e dunque restituiscono un quadro che risale ormai a tre lustri fa. Inoltre

Tabella 5 - *Accettabilità delle forme perifrastiche con tue 'fare' in greschòneytitsch* (Angster 2011: 83; con modifiche). [agramm. = agrammaticale; incomp. = incompatibile; opz. = opzionale; obbl. = obbligatorio]¹⁵

	ausiliari e modali	wéssò 'sapere'	KV	altri verbi lessicali	tóntz causativo
Indicativo presente	agramm.	incomp.	opz.	opz.	obbl.
Imperativo	agramm.	incomp.	opz.	opz.	obbl.
Konjunktiv I	agramm.	incomp.	opz.	opz.	obbl.
Konjunktiv II	agramm.	incomp.	competenza passiva delle forme sintetiche		obbl.

Il test di accettabilità mostra inoltre la resistenza all'uso della *Tun-Periphrase* di wéssò 'sapere'. Questo verbo partecipa ad alcune delle caratteristiche formali dei modali (si veda sopra nota 7), ma dal punto di vista semantico-funzionale non appartiene né alla classe dei modali, né agli ausiliari. La resistenza di wéssò alla perifrasi ha i tratti di una sostanziale incompatibilità e in quanto tale è maggiore di quella dei *Kurzverben*, come visto anche dai dati in (5). I rimanenti verbi lessicali, qualunque sia la loro classe flessiva di appartenenza (verbi forti o deboli di qualunque classe) sono compatibili con la *Tun-Periphrase*, ma, anche per quei verbi che occorrono nei dati sempre in forme analitiche, non si può parlare di obbligatorietà di tali forme e le forme sintetiche prodotte secondo le descrizioni disponibili sono state sempre riconosciute e accettate dai parlanti consultati (Angster 2011: 77).

L'unico caso di obbligatorietà riconosciuta delle perifrasi con *tue* 'fare' è la costruzione causativa, che verrà discussa sotto in § 4.1.

l'informatore più anziano, purtroppo defunto, avrebbe raggiunto nel momento in cui scriviamo i 100 anni. Si può dunque ipotizzare che la situazione odierna sia ulteriormente evoluta in direzione di un uso ancor maggiore delle *Tun-Periphrase*.

¹⁵ La differenza tra "agrammaticale" e "incompatibile" nella tabella è stata basata sulla categoricità del rifiuto espresso dai parlanti nei confronti delle forme proposte nel test di accettabilità: maggiore nel caso di agrammaticale, minore nel caso di incompatibile. Va osservato che i casi di agrammaticalità riguardano verbi coinvolti in costruzioni che esprimono perifrasticamente tratti TAM.

4. *Cambiamento*

Spostando l'attenzione sul cambiamento linguistico, in questo paragrafo si considereranno alcuni sviluppi diacronici che pongono il problema di quanto le innovazioni siano stabili e diffuse nella comunità dei parlanti (e nelle comunità walser a Sud delle Alpi in generale).

La trattazione di questi fenomeni non potrà essere esaustiva e dove possibile si rimanderà per maggiori informazioni a lavori in cui tali fenomeni sono discussi in modo più approfondito. In quanto segue si farà riferimento in particolare alla costruzione causativa nel *titsch* di Gressoney e nelle comunità walser italiane e si toccherà il tema dei pronomi enclitici soggetto e del rinnovamento della morfosintassi verbale.

4.1 Stabilità interna, variabilità diatopica

La perifrasi causativa costruita con l'elemento *tóntz* è un'innovazione del *titsch*. L'appartenenza categoriale di questo elemento in sincronia non è chiara, come vedremo qui sotto, oscillando tra quella di una forma verbale idiosincratca e di un complementatore invariabile. Diacronicamente la sua origine è incerta, ma si può tuttavia concludere con ragionevole sicurezza che sia da cercare tra le forme del verbo *tue* 'fare'.

L'elemento causativo *tóntz* occorre senza modifiche della sua forma sia in frasi al perfetto, sia nella *Tun-Periphrase* in vari tempi e modi, sia retto da modali. Regge l'infinito del verbo indicante l'azione causata (Angster 2011; Angster, Gaeta 2021).

- (6) *Tue de tälloré tóntz wäsche vòn Luis*
 fare.IMP.2SG i piatti CAUS lavare.INF da Luigi
 "Fai lavare i piatti a Luigi" [TESORO: 55]

- (7) *débel häscht du dem Joseph*
 mentre avere.IND. PRS.2SG 2SG.NOM DET.M.DAT.SG Joseph
d'stòré tóntz wéderhole
 la=storia CAUS ripetere.INF
 "mentre tu facevi ripetere a Giuseppe la filastrocca" [TESORO: 62]

Si può vedere come *tóntz* sia retto in (6) dall'imperativo di *tue* 'fare' e a sua volta regga l'infinito *wäsche* 'lavare'. In (7), invece, *tóntz* è retto dall'ausiliare *hä* 'avere' in una frase al *Perfekt*. Nonostante *Tun-Periphrase* e *Perfekt* reggano di norma un infinito e un participio per-

fetto rispettivamente, si può notare come *tòntz* sia immutato nei due contesti morfosintattici.

Stabile a Gressoney (benché in competizione con la perifrasi causativa con *loa* ‘lasciare’, ted. *lassen*) questa costruzione è l’esito di un percorso di grammaticalizzazione i cui stadi iniziali si possono osservare nelle varietà walser circostanti.

A Formazza ad esempio il causativo è espresso sempre da ‘fare’ a reggere un’infinitiva introdotta da *z* (ted. *zu*):

- (8) *Tö dem Luis t blattulti z wäschä*
 fare.IMP.2SG DET.M.DAT.SG luis i piatti COMP lavare.INF
 “Fa’ lavare i piatti a Luigi” [TESORO: 55]
- (9) *hescht [...] z sägä ta*
 avere.IND.PRS.2SG COMP dire.INF fare.PART.PRT
 “facevi (lett. hai fatto) ripetere” [TESORO: 62]

A differenza di quanto avviene nella varietà gressonara, nel *titsch* formazzino la costruzione causativa consiste in una perifrasi con *fare* che regge un’infinitiva invece dell’infinito semplice. Per questa ragione in (8) ‘fare’ occorre all’imperativo (ma lo stesso potrebbe valere per il presente indicativo) senza la necessità dell’intervento della *Tun-Periphrase*, mentre al perfetto l’ausiliare ‘avere’ regge ‘fare’ con funzione causativa al participio (*ta* ‘fatto’). Per ragioni di spazio si rimanda ad Angster (2011) e Angster, Gaeta (2021) per una discussione delle ipotesi su come la particella causativa *tòntz* si origini in *greschòneytitsch*. Sia sufficiente qui osservare il fatto che queste comunità, così vicine geneticamente, sono tuttavia caratterizzate da strutture anche radicalmente differenti. La loro vicinanza genetica e geografica – ma anche la loro somiglianza culturale e in una certa misura storico-sociale – non assicura dunque che queste varietà condividano i loro tratti grammaticali e ciò lascia aperta la possibilità che la divergenza tra esse possa risultare nella loro appartenenza a tipi linguistici differenti.

4.2 Variabilità interna

La variazione interdialeale, tuttavia, stupisce fino a un certo punto, visto che le comunità walser risultano sostanzialmente isolate le une dalle altre (si veda sopra in § 2.1). Più interessante e in certo modo sorprendente è la variazione interna. Per illustrarla si considererà ora il caso dei pronomi enclitici soggetto.

Già Zürrer (1982: 94) attira l'attenzione sullo sviluppo potenziale di nuove desinenze verbali date dalla rianalisi dei pronomi clitici soggetto. Giacalone Ramat (1989: 42) giunge a proiettare questo sviluppo e presenta un possibile paradigma di tali forme innovative¹⁶.

Tabella 6 - *Innovazioni nell'indicativo presente in greschèneytitsch*
(Giacalone Ramat 1989: 42)

<i>Indicativo Presente</i>	
1S	<i>tuen-é goa</i>
2S	<i>tuescht goa</i>
3S	<i>tuet-er (M), tuet-(d)sch (F), tuet-s goa (N)</i>
1P	<i>tie-ber goa</i>
2P	<i>tied-er goa</i>
3P	<i>tien-dsch goa</i>

L'enclisi dei pronomi soggetto non è un fenomeno nuovo o ignoto ad altre varietà tedesche: Bohnenberger (1913) indica sistematicamente per le varietà vallesane diverse serie pronominali, toniche e atone. Un'analoga distinzione inoltre si osserva nelle varietà alemanniche della Svizzera (Russ 1990: 375-76).

La distribuzione delle forme atone – che, oltre ai pronomi al nominativo, includono anche forme di accusativo, dativo e, più di rado, genitivo – segue un modello specifico. Nei dati scritti del *titsch* di Gressoney si possono riconoscere vari esempi di occorrenza dei pronomi enclitici soggetto nei contesti sintattici tipici del *verb-second* che può essere descritto, riprendendo Harbert (2007) come “l'obbligo che il verbo finito di una proposizione non sia posto ad una distanza maggiore dall'inizio della proposizione stessa della seconda posizione [sintattica] (escludendo dal conto le congiunzioni)”¹⁷.

- (10) *Khieme=ber* *äch* *noch* *bi* *tag*
venire.IND.PRS.1PL=1.PL PTCL ancora entro giorno

¹⁶ Si noti come Giacalone Ramat (1989) presenti tali forme innovative associate alla *Tun-Periphrase*. Curioso che *goa* 'andare' sia scelto per illustrare questo fenomeno, visto che, come menzionato in precedenza, è tra i *Kurzverben* più resistenti alla *Tun-Periphrase*.

¹⁷ «Descriptively speaking, the V-2 phenomenon is the requirement, apparently holding under at least some circumstances in all of the GMC languages, that the finite verb of the clause be no further from the beginning of the clause than second position (not counting conjunctions)» (Harbert 2007: 398).

zäm *Hus?*
 a.DET.N.DAT.SG casa(N)

‘Arriveremo ben a casa prima di notte?’ [CAW; *Duezòmoal* 1982]

- (11) *Vor dem prozess hätt=er*
 prima DET.M.DAT.SG processo(M) avere.IND.PRS.3SG=3SG.M
kät en paar schnaps-iene tronget
 avere.PART.PRT un paio grappa-DIM bere.PART.PRT
 ‘Prima del processo aveva bevuto un paio di grappini’
 [CAW; *Duezòmoal* 1995]

Entrambi gli esempi sono pubblicati sul bollettino parrocchiale nella rubrica *Duezòmoal*, curata per più di 20 anni da una coppia di autrici nate prima della Seconda Guerra Mondiale che in essa si impegnavano a raccogliere testimonianze della cultura tradizionale di Gressoney ‘perché non tutto vada dimenticato’ (sottotitolo della rubrica: *fer dass nid alz ganngé ém Vergäs*). Le stesse autrici hanno partecipato alla fondazione del *Centro Culturale Walser/Walser Kulturzentrum* e hanno lavorato alla stesura del vocabolario del *titsch*, riferimento anche normativo per la varietà di Gressoney.

L’esempio (10) contiene un’interrogativa polare e, sebbene in prima posizione non ci sia alcun elemento, il pronome soggetto si sposta in enclisi. L’esempio (11) mostra invece un caso in cui il pronome soggetto va in enclisi perché il *Vorfeld* è occupato¹⁸. Nel corpus *ArchiWals* si può riconoscere come la cliticizzazione dei pronomi soggetto occorra nei contesti visti sopra.

Tuttavia è anche possibile trovare nel corpus scritto *ArchiWals* esempi di cooccorrenza del clitico con un soggetto pronominale o nominale. Questi esempi suggeriscono un’evoluzione dei clitici nella direzione della loro rianalisi come flessione verbale.

- (12) *Khein mòdernitét chann das*
 nessun modernità(F) potere.IND.PRS.3SG DEM.N.NOM.SG
vernéchte wenn wier allé
 distruggere.INF se 1.PL tutti

¹⁸ Il termine *Vorfeld* è usato nella teoria topologica della struttura della frase tedesca (si veda Wöllstein 2014) per indicare la parte di frase che precede il verbo di forma finita posto in seconda posizione. Secondo tale teoria il verbo di forma finita (*V1*) e quello di forma non-finita (*V2*) suddividono la frase in tre campi (*Felder*): *Vorfeld*, prima di *V1*; *Mittelfeld* tra *V1* e *V2* dove si trovano in genere i complementi; *Nachfeld*, dopo *V2*, dove si pongono le eventuali subordinate o talvolta alcuni complementi.

wel=ber=s *erette.*
 volere.IND.PRS.SG=1.SG=3SG.N.ACC salvare.INF
 ‘Nessuna modernità lo può distruggere se noi tutti lo vogliamo
 salvare.’ [CAW; Duezòmoal 1999]

- (13) *Hilde kròa vòr Lagnò tut=z gä*
 Hilde appena prima Alagna fare.3SG=3SG.N dare.inf
da gédangè z'=néntsch tischò.
 DEF.M.ACC.SG idea(M) COMP=1PL.ACC cambiare.INF
 ‘Hilde proprio prima di Alagna dà l’idea di cambiarci’ [CAW: 2007]

È interessante notare come tali esempi, benché rari, occorrono in testi abbastanza recenti (1999 e 2007, rispettivamente) di due parlanti in un certo modo marginali rispetto alla “norma” del *titsch* incarnata dal *Walser Kulturzentrum*: un parlante che da anni vive a Jakarta (12) e una giovane gressonara (13)¹⁹.

Il confronto degli esempi (10)-(13) stimola alcune osservazioni riguardanti la variazione. Nell’ambito dell’evoluzione (non compiuta) dei pronomi enclitici soggetto da forme atone il cui uso è determinato sintatticamente a desinenze innovative della coniugazione verbale, i testi contenuti in *ArchiWals* non mostrano uno stadio particolarmente avanzato su questo percorso di grammaticalizzazione, vuoi perché molti testi sono frutto della penna di pochi autori, vuoi perché esiste una forma di norma linguistica implicita emanata dagli appartenenti alla locale associazione culturale, spesso coincidenti con gli autori dei testi. Deviazioni da questa norma esistono e corrispondono ad autori lontani (per età o residenza) da quel nucleo normativo.

4.3 Variabilità interna e diamesia

Si è osservato sopra come nel mezzo scritto la variazione interna al *titsch* sia legata a variabili sociolinguistiche quali età e rete sociale. Valuteremo ora se e come si attui la variazione sul piano diamesico. Considerando nuovamente l’etnotesto del *PALWaM* riportato in (5) sopra e ripetuto qui sotto in (14) alla luce della cliticizzazione dei pronomi soggetto, possiamo notare come il parlato presenti una situa-

¹⁹ Aneddoticamente, ma a dimostrazione del livello di aderenza alla “norma del *titsch*” del testo prodotto dalla giovane autrice, l’esemplare del bollettino parrocchiale da cui da cui esso è stato raccolto presentava delle correzioni, per lo più ortografiche, operate probabilmente da un parlante più anziano del Centro Culturale Walser (luogo dove i testi dei bollettini sono stati materialmente acquisiti nell’ambito dei progetti *DiWaC* e *ArchiWals*).

zione differente da quella vista nello scritto. Le forme con pronomi enclitici, soggetto e non, sono sottolineate nel brano e seguite da una glossa (il brano è tradotto sopra in nota 10).

(14)

VC: eh, dunque, *du weféll chie häschti?*

L: *hänn=é=ró* (avere.IND.PRS.1SG=1SG=3PL.GEN) *só*
fénföntzwentzg en allem aber z mälche elwé

VC: *elwé jetza*

L: *jetza, ja, ja ón em sómmer es bétzié mé wóróm gam=ber* (andare.
IND.PRS.1PL=1PL) *óf z alpó de*

VC: certo

L: *tie=be=ró* (fare.IND.PRS.1PL=1PL=3PL.GEN) *gé es pare z zueft*
ón ón sir=ró (essere.3PL=3PL.GEN) *es bétzié mé*

VC: *ón em sómmer tuescht óu machó de chésch?*

L: *ja*

VC: *anschtatt jetza es bétz wenégor*

L: *ja, ja ma jetza schier néks fórom sinn=tsch* (essere.3PL=3PL) *d*
chalbiene ón

VC: *ón tien dšchi trénge d mëlch*

L: *de tien dšchi trénge d mëlch bés wenn sinn gmaschté ón de tie=ber*
(fare.IND.PRS.1PL=1PL), *de tuen=é* (fare.IND.PRS.1SG=1SG)
de afoa z chéschó

VC: *ón wewéll moal z tagsch tuescht mälché?*

L: *zwei moal z tagsch de morgé ón em oabe*

VC: *ón wewéll mëlch? Tientsch* (fare.3PL=3PL) *machó?*

L: *ma óngéfer d chue wie*

VC: certo *fón*

L: *fón dri litter só bés zwelfé, dritzené.*

Nel brano, come già detto in precedenza, sono presenti 18 turni e 14 forme verbali finite. Tra esse si hanno tre occorrenze di pronome soggetto tonico e 8 occorrenze di pronome soggetto enclitico. Nei casi di uso del soggetto enclitico tre occorrenze vanno contro le regole del *verb-second*: due occorrenze sono a inizio frase e una segue la congiunzione *ón* 'è'. Le restanti 5 rispettano invece le regole del *verb-second*: due seguono l'avverbio *de* 'allora'; due seguono l'introduttore di subordinata *wóróm/fórom* 'perché'. Si ha inoltre un caso in cui il verbo è contenuto in una frase interrogativa. In quattro casi il verbo non ha alcun soggetto esplicito, né nominale, né pronominale: in due casi si tratta di forme del verbo *si* 'essere' e in due casi si tratta di verbi alla se-

conda persona singolare (unico pronome personale a non presentare forma atona; si veda anche sopra Tabella 4)²⁰.

Non c'è qui lo spazio di analizzare ulteriormente i fenomeni fonologici di assimilazione al confine tra desinenze e clitici e tra clitici. Si noti soltanto come le forme di terza plurale genitivo *-rò* tendano a generare un'assimilazione regressiva che oscura l'esito delle desinenze verbali, si veda in particolare *sir=rò* 'essere.3PL=3PL.GEN', in cui la forma verbale sarebbe *sinn* 'essere.3pl'²¹. Analogamente il clitico soggetto di prima plurale *-ber* (con forma tonica *wier*, ted. *wir*), ad esempio in *tie=ber* 'fare.1PL=1PL' causa l'assimilazione della nasale finale della forma verbale *tien* 'fare.1PL'.

Pur minoritari (3 su 8), i casi di cliticizzazione al di fuori dei contesti tipici del *verb-second* – e cioè a inizio di frase affermativa – mostrano come nel parlato l'occorrenza dei pronomi soggetto in enclisi sia più ampia di quanto non avvenga nello scritto.

Questo tratto, così come il frequente uso della *Tun-Periphrase* visto sopra in § 3.3 suggerisce una differenza tra il corpus di testi contenuti in *ArchiWals* – in gran parte prodotti tra gli anni '70 e gli anni 2000 – e i dati di parlato pubblicati nel *PALWaM* – raccolti nei primi anni '10.

Tale differenza si può ascrivere a diverse dimensioni di variazione sociolinguistica: innanzitutto il diverso mezzo usato (diamesia); oppure la diversa età dei parlanti/scriventi (*age-grading/ apparent time*, v. Labov 1994: 45-46); è anche possibile ipotizzare l'esistenza di una (seppur debole) norma del *titsch*, specialmente allo scritto (diafasia)²².

²⁰ Si noti che le desinenze di seconda persona singolare del tedesco si possono analizzare diacronicamente proprio come l'univerbazione tra la forma di seconda persona singolare con desinenza *-s* e un pronome soggetto enclitico *thu/du*. Forme univerbate attestate precocemente come *forsabbistu, gilaubistu*, mantengono la vocale del pronome che più tardi scompare. L'innovazione appare nel IX secolo in francone e nel X secolo in tedesco superiore (Braune & Eggers 1987: 258). Altre aree dialettali, in particolare il tedesco medio occidentale, conservano invece più a lungo, fino al periodo alto tedesco medio, le desinenze in *-s* (Paul 2007: 243).

²¹ Per il vocalismo delle forme atone del pronome di terza persona plurale genitivo si confronti la forma alto tedesco antica *iro*, accanto a *ira* e *iru* (Braune & Eggers 1987: 239, 241).

²² Un revisore anonimo fa notare che una dinamica simile è stata osservata per il piemontese in Ricca (2008). Nelle corpus scritto giornalistico analizzato dall'autore l'aderenza al sistema di clitici proposto da alcune descrizioni della koiné normativa (ad

La rilevanza di ciascuna di tali dimensioni di variazione andrebbe vagliata attentamente su di un inventario di fenomeni e su di un campione di dati – in particolare di parlato – più ampio.

Per ciò che riguarda lo scritto, nonostante le difficoltà note legate all'uso dei corpora scritti per l'analisi sociolinguistica (Romaine 2009: 110), per il *titsch* di Gressoney un'analisi sociolinguistica piuttosto precisa è almeno in parte possibile grazie al fatto che gli autori dei testi raccolti in *ArchiWals* sono per lo più noti e se ne può ricostruire un profilo sociolinguistico abbastanza accurato.

5. *Discussione*

Al di là dei limiti dell'analisi qui proposta possiamo notare come nel *titsch* di Gressoney si riconoscano un buon numero di fenomeni soggetti a variazione rispetto alle varietà circostanti strettamente affini, lungo le dimensioni di variazione sociolinguistica e lungo una diacronia "compressa" dai radicali cambiamenti sociali e culturali sperimentati dalla comunità dei parlanti negli ultimi decenni.

I cambiamenti evidenziati per la morfologia verbale mostrano da un lato la sostanziale preservazione delle differenziazioni interne al paradigma verbale del *titsch*, ma non senza conseguenze tipologiche.

La morfologia verbale del *titsch* infatti sembra evolvere in senso analitico. Tale evoluzione pare anche confermata dalla grammaticalizzazione della particella causativa *tòntz*, unico contesto di stretta obbligatorietà della *Tun-Periphrase*.

La (incompiuta) rianalisi dei pronomi atoni enclitici come desinenze verbali, invece, mostra l'interazione tra l'evoluzione morfologica e quella sintattica. La sequenza *forma verbale finita-forma pronominale atona* si ottiene inizialmente in conseguenza del fenomeno del *verb-second*, come attestato nel corpus scritto. Tuttavia, se si considerano le produzioni orali più recenti, le regolarità riscontrate nei testi scritti vengono meno.

Se consideriamo i fenomeni evidenziati sopra nei termini del rapporto tra semplificazione e complessificazione, possiamo osservare tendenze in direzioni opposte.

es. Brero & Bertodatti 1988 e Villata 1997) è più stringente, mentre è disattesa nel corpus di parlato analizzato tratto da Bonato (2004) (Ricca 2008: 126).

La svolta in senso analitico del sistema verbale del *titsch* costituisce senz'altro un caso di semplificazione, almeno a livello morfologico, visto che riduce la varietà di forme verbali sintetiche da apprendere a quelle di pochi verbi con funzioni grammaticali e limita la rilevanza delle classi flessive verbali al solo contrasto tra infinito e participio perfetto. Tale processo di semplificazione, tuttavia, non incide, come visto, sull'integrità del sistema di distinzioni all'interno dei paradigmi. Inoltre va considerato che lo sviluppo cruciale è il dilagare della *Tun-Periphrase*, fenomeno che non appare innovativo qualitativamente – esiste come visto da tempi remoti –, ma solo quantitativamente. In questo senso non sembra che il recente successo della *Tun-Periphrase* si debba ascrivere ad un aumento del contatto con l'area romanza, ma può essere spiegato anche come conseguenza della riduzione del contatto con l'area tedesca e del venir meno di una pressione normativa del tedesco letterario (in parte persistente nella norma locale rappresentata dall'associazione culturale).

In direzione diametralmente opposta va invece il fenomeno della cliticizzazione, visto che, pur nel contesto del consolidarsi della *Tun-Periphrase*, essa tende ad innovare il sistema verbale nuovamente in senso sintetico – almeno per ciò che riguarda la marcatura personale – nonché alla creazione di gruppi clitici – come quelli visti sopra in 4.3 – di notevole complessità sia dal punto di vista morf fonologico, sia dal punto di vista della cumolazione dell'informazione grammaticale. Questo fenomeno, al confine tra sintassi e morfologia, può essere interpretato come un processo di complessificazione e apparterebbe a quei fenomeni linguistici definiti “maturi” (*mature phenomena*, Dahl 2004) per il cui sviluppo sono necessari periodi di tempo lunghi e condizioni che con minor probabilità si presentano in comunità ampie, soggette a elevato contatto, instabili e sprovviste di reti sociali compatte (Trudgill 2011: 149).

Tutti i fenomeni presi in considerazione mostrano comunque come gli stadi evolutivi attestati correlino con le dimensioni di variazione sociolinguistica e cioè:

- con gli sviluppi storici, sociali e culturali dell'ultimo secolo;
- con canali di comunicazione differenti (parlato vs. scritto);
- con le scelte di diversi autori più o meno vicini al centro normativo rappresentato dall'associazione culturale presso la quale sono stati nel tempo prodotti e pubblicati la gran parte dei testi.

5.1 Greschèneytitsch: quale fascio di varianti?

Pur nel mantenimento del sistema di distinzioni funzionali all'interno del sistema linguistico del *titsch*, pochi tra i vari fenomeni considerati appaiono stabili. Immaginando allora che il *titsch* sia una varietà linguistica sconosciuta e poco documentata nonché genealogicamente esotica, si potrebbe porre il problema di come completare un suo profilo tipologico.

Limitandoci ai fenomeni considerati, la tendenza analitica data dalla *Tun-Periphrase* è molto marcata, ma non è un tratto obbligatorio se non nella nicchia del causativo. La configurazione sintattica e il comportamento dei clitici, invece, sono fenomeni soggetti ad una elevata variazione interna.

Quale fascio di varianti andrebbe dunque privilegiato in una descrizione? Quali varianti andrebbero scelte per caratterizzare il tipo linguistico del *titsch* di Gressoney?

Almeno due sono le scelte possibili. La prima sarebbe quella di optare per la varietà scritta, che ha tendenze normative, è dotata di un certo riconoscimento nella comunità, ma che dai dati dell'*ArchiWals* appare comunque come il risultato delle produzioni di pochi autori, forse anche influenzati dall'onda lunga dell'influenza normativa del tedesco letterario, riferimento culturale e linguistico per la comunità di Gressoney almeno fino all'inizio del secolo scorso.

La seconda possibilità invece è quella di privilegiare la varietà orale, che però appare instabile e in via di scomparsa – la maggior parte degli autori dei testi inclusi in *ArchiWals* sono morti, i parlanti con piena competenza si riducono anno dopo anno – e che è evidentemente più avanzata nei processi di grammaticalizzazione, forse anche per il sopraggiunto isolamento dall'area tedescofona e i sempre più intensi contatti con l'ambito romanzo a partire dalla fine della Seconda Guerra Mondiale.

6. Conclusione

Indagare le varietà walser e in particolare il *titsch* di Gressoney ha fatto emergere un quadro di variazione complesso e frastagliato che risulta in parte sorprendente se si considera che il numero di parlanti di questa varietà è estremamente ridotto e che anche in passato non è mai stato superiore alle poche centinaia.

Il contatto linguistico, che in questo contributo abbiamo inteso non tanto dal punto di vista della penetrazione di tratti esogeni, quanto da quello delle conseguenze sugli sviluppi linguistici interni delle mutate condizioni sociali, linguistiche e culturali – in primo luogo l'ingresso dell'italiano a spese del tedesco come lingua di istruzione prima e socializzazione poi – ha evidenti conseguenze nell'accelerare il cambiamento linguistico.

Si può ipotizzare nel caso del *titsch* di Gressoney uno sviluppo simile a quello osservato da Dorian (1994) per il villaggio di Embo: dopo l'interruzione dei contatti con l'area tedescofona a inizio '900, i parlanti si sono trovati gradualmente esposti all'italiano, ma in un quadro sociale di persistente compattezza e di relativo isolamento grazie al quale la comunità di parlanti non si è immediatamente smembrata, ma ha ancora per un certo tempo assimilato gli elementi provenienti dall'esterno. Il venir meno della norma del tedesco e il sopraggiunto contatto hanno però accelerato il cambiamento linguistico. Un cambiamento linguistico accelerato crea variazione interna e instabilità del sistema linguistico.

Riconsiderando alla luce dei dati del *titsch* analizzati la posizione di Gressoney nelle categorie proposte da Trudgill (2011; si veda sopra § 2.3.) possiamo considerare i fenomeni di complessificazione compatibili con l'evoluzione della comunità dalla categoria 1 a quella 2 con l'aumento del contatto nel contesto di una comunità socialmente compatta. Nel *titsch* prosegue infatti la maturazione della complessa morfosintassi verbale legata ai cluster di pronomi enclitici. La situazione odierna – che corrisponde alla categoria 4 per via del crollo della stretta rete sociale persistente fino almeno alla metà degli anni '60 del '900 – non è stata ancora analizzata in modo soddisfacente. Il passaggio a questa nuova fase però potrebbe essere compatibile con l'ampliamento della *Tun-Periphrase* ai *Kurzverben* e con l'aumento della sua frequenza d'uso: questo cambiamento contribuisce infatti a semplificare – sebbene solo in parte – il sistema di classi flessive verbali. Promettente dal punto di vista dell'identificazione dei fenomeni di cambiamento più recenti e ipoteticamente legati all'allentamento della rete sociale sarebbe l'analisi del parlato, in particolare dei parlanti più giovani e di coloro che si pongono ai margini della seppur implicita "norma" del *titsch*.

Tornando al tema iniziale del campionamento linguistico, la discussione qui condotta suggerisce alcune riflessioni. In primo luogo che il contatto linguistico non può essere liquidato semplicemente come un fenomeno regolare che riguarda ogni lingua, come suggerisce Bakker (2011: 6). Le lingue sottoposte a intenso contatto possono presentare caratteristiche simili, ad esempio una tendenza alla semplificazione (come suggerisce Trudgill 2011) e come si può osservare anche dai dati del *titsch*, se si interpreta in questo modo il dilagare della *Tun-Periphrase*. In secondo luogo, alla luce di quanto considerato, non si può nemmeno dare per scontato che una lingua piccola – in quanto rappresentante di una società piccola – sia necessariamente meno complessa come invece suggerisce Grandi (2019). Il caso del *titsch* mostra infatti, come anche questa piccola comunità con ormai pochi parlanti esprima un diasistema che si sviluppa su diverse dimensioni della variazione sociolinguistica, ciascuna caratterizzata da specifiche varianti. Infine, l'esempio del *titsch* suggerisce anche che, a prescindere dalle dimensioni della società considerata la presenza (o l'emergere) di una norma scritta costituisce un fattore da non sottovalutare nel quadro della stabilità di una varietà linguistica.

Le riflessioni qui condotte mostrano da un lato la rilevanza di un profilo sociolinguistico di una lingua nel contesto del campionamento tipologico. Conoscere e analizzare il caso di varietà come il *titsch* – piccole e in decadenza, ma al tempo stesso “note” dal punto di vista della ricca rete di informazioni linguistiche e storiche sui loro parenti stretti e su loro stesse – può permettere di approfondire la conoscenza delle dinamiche di una tipologia sociolinguistica come quella proposta da Trudgill (2011). Le osservazioni così raccolte, se generalizzabili in modo convincente, possono essere poi applicate all'analisi del caso di altre lingue per le quali il bagaglio di conoscenze a disposizione è invece molto più limitato. In questo modo una tipologia sociolinguistica matura potrebbe arricchire come ulteriore fattore il campionamento tipologico e portare a generalizzazioni che abbraccino anche la variazione sociolinguistica.

Dall'altro lato il caso della variazione interna al *titsch* è esemplare anche dal punto di vista della documentazione linguistica (ed eventualmente anche della pianificazione linguistica). Una varietà, anche se parlata da poche decine di persone, può presentare diversi fasci di tratti mostrando quanto possa essere delicato il compito del linguista

impegnato nella documentazione o di quello impegnato nella pianificazione di una varietà di riferimento per azioni di politica linguistica.

Ringraziamenti

Il presente contributo nasce dalle riflessioni accumulate dall'autore negli anni partecipando a diversi progetti di ricerca e considerando vari aspetti linguistici e sociolinguistici delle varietà walser, in particolare del *titsch* di Gressoney. Ringrazio quanti hanno partecipato a questo percorso, siano stati informatori, colleghi, coautori.

Ringrazio i due revisori anonimi per i loro preziosi commenti sulla prima versione del manoscritto e i curatori di questo volume.

Riferimenti bibliografici

- Angster, Marco. 2004-05. *La perifrasi tue + infinito nel titsch di Gressoney*. Università di Torino. (Tesi di laurea triennale.)
- Angster, Marco. 2011. Il verbo fare a Gressoney. Caratteri e forme di forme verbali analitiche in un dialetto walser. In Fazzini, Elisabetta (a cura di), *Il tedesco superiore. Tradizione scritta e varietà parlate*, 65-86. Alessandria: Dell'Orso.
- Angster, Marco. 2012. Isolamento e contatto. Stratigrafia del lessico dei walser meridionali dai dati del PALWaM. *Bollettino dell'Atlante Linguistico Italiano (III Serie)* 36. 155-200.
- Angster, Marco. 2014. Lingue di minoranza e di maggioranza. 200 anni di lingue straniere a Gressoney (AO). In Porcellana, Valentina & Diémoz, Federica (a cura di), *Minoranze in mutamento: Etnicità, lingue e processi etnografici nelle valli alpine italiane*, 105-121. Alessandria: Dell'Orso.
- Angster, Marco & Dal Negro, Silvia. 2017. Linguistische Distanz einschätzen: Der Fall von Walserdeutsch im Licht von Lexikalischen Daten und Soziolinguistischen Parametern. In Christen, Helen & Gilles, Peter & Purschke, Christoph (a cura di), *Räume, Grenzen, Übergänge. Akten des 5. Kongresses der Internationalen Ghesellschaft für Dialektologie des Deutschen (IGDD)*, 9-25. Stuttgart: Franz Steiner Verlag.
- Angster, Marco & Gaeta, Livio. 2018. Wie kurz sind die Kurzverben? Morphologische Merkmale in Gressoney und Issime. In Rabanus, Stefan (a cura di), *Deutsch als Minderheitensprache in Italien. Theorie und*

- Empirie kontaktinduzierten Sprachwandels, Germanistische Linguistik*, 239-240. 211-237.
- Angster, Marco & Gaeta, Livio. 2021. Contact phenomena in the verbal complex: the Walser connection in the Alpine area. *STUF - Language Typology and Universals / Sprachtypologie und Universalienforschung*. 74(1), 73-107.
- Angster, Marco & Bellante, Marco & Cioffi, Raffaele & Gaeta, Livio. 2017. 'I progetti DiWaC e ArchiWals'. In Livio Gaeta (a cura di), *Le isole linguistiche tedesfone in Italia: situazione attuale e prospettive future* (Workshop, Torino, 24 febbraio 2017), *Bollettino dell'Atlante Linguistico Italiano*, 41. 83-94.
- Antonietti, Federica & Valenti, Monica & Angster, Marco (a cura di). 2015. *Piccolo atlante linguistico dei Walser meridionali*, Aosta, Tipografia Valdostana.
- Bakker, Dik. 2011. Language Sampling. In Song, Jae Jung (a cura di), *The Oxford Handbook of Linguistic Typology*, 1-19. Oxford: Oxford University Press (Oxford Handbooks Online).
- Bohnenberger, Karl. 1913. *Die Mundart der deutschen Walliser im Heimattal und in den Aussenorten*. Huber Verlag: Frauenfeld. (Beiträge zur schweizerdeutschen Grammatik. VI).
- Bonato, Massimo. 2004. *Tratti variabili nella sintassi del piemontese parlato contemporaneo*. Università di Torino. (Tesi di laurea).
- Braune, Wilhelm & Eggers, Hans. 1987. *Althochdeutsche Grammatik / von Wilhelm Braune. Bearbeitet von Hans Eggers. 14. Auflage*. Tübingen: Niemeyer.
- Brero, Camillo & Bertodatti, Remo. 1988. *Grammatica della lingua piemontese. Parola – vita – letteratura*. Torino: Piemont-Europa.
- Campbell, Lyle. 1985. Areal Linguistics and its implications for Historical Linguistics. In Fisiak, Jacek (a cura di), *Papers from the VIIth International Conference on Historical Linguistics, Poznań. 22-26 August 1983*, 25-56. Amsterdam / Poznań: John Benjamins / Adam Mickiewicz University Press.
- Campbell, Lyle. 2006. Areal Linguistics: A Closer Scrutiny. In Matras, Yaron & McMahon April & Vincent Nigel (a cura di), *Linguistic Areas. Convergence in Historical an Typological Perspective*, 1-31 New York: Palgrave McMillan.
- CELE 2002. = CELE (Centre d'Études Linguistiques pour l'Europe). 2002. *Plurilinguisme administratif et scolaire en Vallée d'Aoste. Plurilinguismo amministrativo e scolastico in Valle d'Aosta*. policopiato.

- Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity*. Amsterdam: Benjamins.
- Dal Negro, Silvia. 2004. *The decay of a language. The case of a German dialect in the Italian Alps*. Bern, New York: Peter Lang.
- Dorian, Nancy. 1973. Grammatical Change in a Dying Dialect. *Language* 49(2). 413-438.
- Dorian, Nancy. 1994. Varieties of Variation in a Very Small Place: Social Homogeneity, Prestige Norms, and Linguistic Variation. *Language* 70(4). 631-696.
- Eufe, Rembert & Mader, Anna. 2018. 'Das Walserdeutsche im deutschen und italienischen Sprachgebiet', in: Eller-Wildfeuer, Nicole & Rössler, Paul & Wildfeuer, Alfred (a cura di). *Alpindeutsch. Einfluss und Verwendung des Deutschen im alpinen Raum*, 113-139. Regensburg: Vulpes.
- Fischer, Anna. 2018. *Präteritumschwund im Deutschen. Dokumentation und Erklärung eines Verdrängungsprozesses*. De Gruyter.
- Gaeta, Livio. 2018. Im Passiv sprechen in den Alpen. *Sprachwissenschaft* 43.2. 221-250.
- Geyer, Ingeborg & Angster, Marco & Benedetti, Marcella (a cura di). 2014. *Il tesoro linguistico delle isole germaniche in Italia / Wortschatz aus dem deutschen Sprachinseln in Italien*. Luserna.
- Giacalone Ramat, Anna. 1979. *Lingua, dialetto e comportamento linguistico. La situazione di Gressoney*. Aosta: Musumeci.
- Grandi, Nicola. 2019. Su alcune possibili interazioni tra tipologia e sociolinguistica. *CLUB Working Papers in Linguistics* 3, 2019, 257-265.
- Harbert, Wayne. 2007. *The Germanic Languages*. Cambridge: Cambridge University Press.
- Hill, Eugen. 2010. A case study in grammaticalized inflectional morphology. Origin and development of the Germanic weak preterite. *Diachronica* 27:3 (2010). 411-458.
- Labov, William. 1994. *Principles of Linguistic Change. Volume 1: Internal Factors*. Oxford, UK: Blackwell.
- Langer, Nils. 2000. Zur Verbreitung der tun-Periphrase im Frühneuhochdeutschen. *Zeitschrift für Dialektologie und Linguistik*, 67/3. 287-316.
- Musso, Michele. 2017. Comunità alemanne e francoprovenzali: il quadro toponomastico nella media Valle del Lys. In Musso, Michele (a cura di). *L'espressione linguistica dello spazio in un'area plurilingue: il paesaggio toponomastico della media Valle del Lys*, 9-16. Aosta: Tipografia Valdostana.

- Nübling, Damaris. 1997. Der alemannische Konjunktiv II zwischen Morphologie und Syntax: Zur Neuordnung des Konjunktivsystems nach dem Präteritumschwund. In: Ruoff, Arno und Löffelad, Peter (a cura di). *Syntax und Stilistik der Alltagssprache: Beiträge der 12. Arbeitstagung zur alemannischen Dialektologie, 25. bis 29. September in Ellwangen, Jagst*, 107-121. Tübingen: Niemeyer.
- Paul, Hermann. 2007. *Mittelhochdeutsche Grammatik. 25. Auflage*. Tübingen: Niemeyer.
- Ricca, Davide. 2008. Tratti instabili nella sintassi del piemontese contemporaneo: tra italianizzazione e arcaismi locali. In Heimann, Sabine (a cura di). *Sprachwandel und (Dis-)Kontinuität in der Romania*, 113-127. Tübingen: Niemeyer.
- Rizzi, Enrico. 2002. Le fonti. In Luigi Zanzi & Enrico Rizzi (a cura di), *I Walser nella storia delle Alpi: un modello di civilizzazione e i suoi problemi metodologici. Seconda edizione*. 441-516. Milano: Jaca Book.
- Romaine, Suzanne. 2009. Corpus linguistics and sociolinguistics. In Lüdeling, Anke & Kytö, Merja (a cura di). *Corpus Linguistics. An International Handbook. Volume 1*, 96-111. Berlin, New York: Walter de Gruyter.
- Russ, Charles V. J. 1990. High Alemannic. In Russ, Charles V. J. (a cura di). *The dialects of modern German*, 364-393. London: Routledge.
- Schott, Albert. 1842. *Die deutschen colonien in Piemont. Ihr land, ihre mundart und herkunft. Ein beitrug zur geschichte der Alpen*. Stuttgart, Tübingen: J. G. Gotta'scher Verlag.
- Squinabol, Barbara. 2008. *Dinamiche linguistiche nella comunità walser di Gressoney*. Università di Pavia. (Tesi di laurea magistrale.)
- Trudgill, Peter. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.
- van der Auwera, Johan. 2011. 15 – Standard Average European. In Kortmann, Bernd & van der Auwera, Johan (a cura di). *The Languages and Linguistics of Europe. A comprehensive guide*, 291-306. Berlin, Boston: de Gruyter Mouton.
- Villata, Bruno. 1997. *La lingua piemontese. Fonologia Morfologia Sintassi Formazione delle parole*. Montréal: L'òsna & Tron.
- WALS = Dryer, Matthew S. & Haspelmath, Martin (a cura di) 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (<http://wals.info>) (Consultato il 09.03.2021.)

- WKZ (1988) = Walser Kulturzentrum. 1988. *Greschèneytitsch. Vocabolario Italiano – Titsch*. Musumeci: Quart (Aosta).
- Wöllstein, Angelika. 2014. *Topologisches Satzmodell*. Heidelberg: Winter.
- Zürrer, Peter. 1982. *Wörterbuch der Mundart von Gressoney. Mit einer Einführung in die Sprachsituation und einem grammatischen Abriss. Frauenfeld*. (Beiträge zur schweizerdeutschen Mundartforschung. XXIV).
- Zürrer, Peter. 2009. *Dialetti walser in contesto plurilingue: Gressoney e Issime in Valle d'Aosta*. Alessandria: Edizioni dell'Orso. [trad. di: Zürrer, Peter. 2009. *Sprachkontakt im Walser Dialekten. Gressoney und Issime im Aostatal (Italien)*. Wiesbaden: Franz Steiner.]

FRANCESCA DI GARBO, ERI KASHIMA,
RICARDO NAPOLEÃO DE SOUZA, KAIUS SINNEMÄKI¹

Concepts and methods for integrating language typology and sociolinguistics

Questo articolo presenta le componenti costitutive di un programma di ricerca per lo studio tipologico dell'adattamento linguistico, ovvero del modo in cui le lingue cambiano in relazione ai contesti socio-storici e ambientali in cui sono utilizzate. Illustriamo una batteria di concetti e metodi volti a comparare sistematicamente contesti sociolinguistici e strutture linguistiche attraverso lo studio di comunità in contatto. Dimostriamo che questi concetti e metodi possono essere usati per studiare i correlati sociolinguistici della diversità e del mutamento linguistico in almeno tre modi: (1) per comprendere le cause del cambiamento linguistico, (2) per creare una base di dati rappresentativi di comunità, fattori sociolinguistici e linguistici, (3) per formulare generalizzazioni sulla base di studi comparativi a livello interculturale e interlinguistico.

This paper presents the building blocks of a comprehensive framework for the typological study of linguistic adaptation, i.e. how languages change in relation to the socio-historical and environmental contexts in which they are used. We showcase a battery of concepts and methods that are geared towards systematically comparing sociolinguistic environments and linguistic structures through the study of communities in social contact. We show that these concepts and methods can be used to investigate sociolinguistic correlates of linguistic diversity and language change in at least three ways: (1) to unravel causal factors related to language change, (2) to create datasets simultaneously addressing selection of communities, sociolinguistic features, and linguistic features, and (3) to formulate generalizations from empirically-grounded cross-cultural and cross-linguistic comparisons.

Parole chiave: adattamento linguistico, contatto linguistico, mutamento linguistico, tipologia, sociolinguistica comparativa

Keywords: linguistic adaptation, language contact, language change, typology, comparative sociolinguistics

¹ All authors equally contributed to this work. The ordering of authors is merely alphabetical.

1. *Introduction*

Research on the non-linguistic correlates of linguistic diversity has been on the rise over the past few decades. Studies from a variety of interconnected fields – such as language typology, sociolinguistics, and language evolution – have shown how patterns of language structures may change under the influence of the larger sociohistorical and environmental contexts in which languages are (or have been) used (see, among others, Wray & Grace 2007; Lupyan & Dale 2010; Trudgill 2011; Bentz & Winter 2013; Everett *et al.* 2015; Sinnemäki & Di Garbo 2018; Blasi *et al.* 2019). These processes of change are loosely captured under the umbrella term *linguistic adaptation* (Lupyan & Dale 2016).

Past studies on linguistic adaptation have individually tackled a diverse range of linguistic and non-linguistic features (e.g. phonemic inventory, morphological complexity for the former; population size, proportions of second language users, and climate for the latter), and investigated how these interact in processes of language change. The findings from these studies, as well as from related areas of research within linguistics, are indicating the need for a more holistic approach (cf. Hruschka *et al.* 2009). The time is thus ripe for establishing a common ground for the systematic study of linguistic adaptation while developing methodologies and tools that are specifically geared towards understand its impact on linguistic diversity and language evolution.

In this paper, we define linguistic adaptation as the processes whereby languages change in a way that enhances their learnability, efficiency of comprehension, and transmissibility in relation to the specific socio-historical and environmental contexts in which they are used. Communicative needs may differ across cultures, depending e.g. on social network structure and density or the amount of language contact. In the context of language contact, we can say that languages may adapt to being learned and used by a bi- or multilingual population. This definition of linguistic adaptation fits with the notion of adaptive changes in cultural evolution, which are defined as those that improve the transmissibility and frequency of a cultural trait (Lupyan & Dale 2016: 650).

Our aim is to contribute to advance research on linguistic adaptation by presenting the research framework that we developed

in the context of the ERC-funded project Linguistic Adaptation: Typological and Sociolinguistic Perspectives to Language Variation (GramAdapt, PI Kaius Sinnemäki). This framework is geared towards the task of comparing sociolinguistic environments and linguistic structures with one another through the in-depth study of communities in social contact. We thus approach linguistic adaptation in a narrow sense, that is by considering changes in language structures that are arguably the result of language contact and the influence of the wider sociolinguistic environment. While we do not assume that all instances of language change in contact situations are by default the result of adaptation processes, through the proposed research design we aim at providing a baseline against which hypotheses about linguistic adaptation can be tested in systematic and controlled ways. We developed the concepts and methods presented in this paper concurrently. The research design is articulated in five parts, each of which we explain below.

Firstly, we conducted a review of proposed explanatory factors for language change under contact situations since our goal is to better understand linguistic adaptation from the perspective of social contact between communities. In this literature review, we primarily consider patterns of socially motivated language variation and change that hinge on some degree of multilingual and bilingual language use because these are the defining features of language contact.

Secondly, we develop a typological approach for comparing sociolinguistic scenarios with one another. Because our goal is to study the relationship between languages and their sociolinguistic environments from a global perspective, we cannot always rely on naturalistic data. The main reason for this is poor availability of language corpora annotated for both linguistic and sociolinguistic features. Instead, our method is based on comparative tools devised by researchers, along the same lines as in state-of-the-art research in language typology.

Thirdly, we develop new methods for selecting sample languages and communities in contact from all around the world because we want to be able to compare linguistic structures and language contact scenarios on a global scale.

Fourthly, we develop a sociolinguistic questionnaire for collecting data on language contact scenarios. The questionnaire design is informed by established knowledge on the dynamics of language use

and language change in speech communities, by our sampling strategy, and by the conceptual tools we developed for comparing sociolinguistic contexts with one another. The sociolinguistic questionnaire is at the center of the data collection process implemented by the GramAdapt project.

The fifth and final part concerns the linguistic variables that we will use in order to test hypotheses about linguistic adaptation. These cover a range of domains of language structures spanning phonology, morphosyntax, and the lexicon. The languages of the sample will be coded for these variables, and their distribution will be then cross-checked with the sociolinguistic profiles emerging from the questionnaire data.

The different parts of the research design relate to each other, in such a way that choices in one subpart (e.g. explanatory factors) affect choices in the other subparts (e.g. questionnaire design). In this paper, we present each of the five parts of this research design starting with explanatory factors for contact-induced change (Section 2), comparative approaches to sociolinguistic environments (Section 3), and language sampling techniques for the investigation of contact scenarios (Section 4). We then illustrate the design principles and workings of the sociolinguistic questionnaire (Section 5) and an overview of the linguistic variables of choice (Section 6). Some concluding remarks are given in Section 7.

2. Explanatory factors for contact-induced change

One straightforward way to establish a common ground for studying linguistic adaptation is to explore the literature that addresses the topic. Given that language phenomena are intrinsically connected to other aspects of human activity, we set out to conduct a literature review on contact-induced change from a broad perspective. That is, instead of focusing on the contact literature alone, in this ongoing review we look at studies from several research domains, including psycholinguistics, bi-/multilingualism, second language learning, language acquisition, and areal linguistics.

In our review, we sought to uncover the more general patterns that may underlie contact-induced change. The choice of the specific areas mentioned above relates to existing mechanisms that have been pro-

posed in the literature to explain contact-induced change. We group these explanations here under the umbrella term *explanatory factors*.

Some of the individual explanatory factors that we identified in the first phase of our research include:

- Openness of the community (e.g. Dahl 2004; Wray & Grace 2007; Trudgill 2011)
- Geographic spread (e.g. Nichols 1992; Atkinson 2011)
- Population size (e.g. Hay & Bauer 2007)
- Number of linguistic neighbors (e.g. Lupyán & Dale 2010)
- Proportion of second-language (L2) speakers in a community (e.g. Bentz & Winter 2013)

Despite these many proposals, the overall picture of how linguistic structures change and diffuse in a community remains fragmentary. Most proposals usually explore a single aspect of the phenomenon without necessarily addressing the more fundamental questions behind the dynamics of contact-induced change. Many of the factors proposed in the literature are in fact proxies for phenomena that lead to structural change, and it remains unclear what their function when tested empirically is.

Some of the questions that the GramAdapt team addresses in the ongoing literature review are:

1. What are the cognitive mechanisms behind contact-induced change?
2. Which mechanisms apply to which changes?
3. What is the influence of social structure on linguistic structure?
4. Which types of social structure contribute to which changes?

More generally, this literature review focuses on gathering the various elements that have been proposed to explain contact-induced change. In order to achieve this, we established a principled way to conduct the literature review. For every source reviewed, we divided our task into two components: the *information extraction component*, and the *evaluation component*. The information extraction component was simply the search for specific pieces of information that relate to our topic of research. The evaluation component involved our own assessment of the various explanatory factors put forward in the literature reviewed. This assessment was also carried out based on a number of criteria that closely suit our research program. In other words, the assessment component aimed at determining how suitable a source was

to our own objectives, rather than evaluating the validity or impact of the materials surveyed. Table 1 summarizes how we structured the literature review through the use of questions.

By systematizing how we approach the vast literature consulted, we were able to maximize the amount of information derived from each source without losing the overall focus. This was especially relevant given the sheer level of detail in each proposal. At the same time, the system we developed allowed us to draw similarities between the various factors in a straightforward way. While this multiprong approach is perhaps not the most suitable to analyze the fine-grained details in specific change processes, it nevertheless allowed us to draw a broad picture of some of the crucial elements behind linguistic adaptation. In the next section we describe some preliminary results of the ongoing review.

Table 1 - *Topics, search components and questions used to structure the literature review*

<i>Task</i>	<i>Identification</i>	<i>Description</i>	<i>Empirical Support</i>	<i>Replicability/Specificity</i>
<i>Information extraction</i>	What name do authors give to this factor?	In which ways does this proposed explanatory factor operate?	What type of data do authors provide to support their proposal?	To which linguistic domain is this proposal applicable? Have authors tested it?
<i>Assessment</i>	To which other proposals can this factor be related?	Are the details of the factor explained in a way that is compatible with language change processes?	How methodologically sound are the data presented? What is the type of sampling technique utilized?	To what extent can this proposal be tested using different data? To what extent can it be used to describe a different linguistic domain?

2.1 Data types

One expected finding given the breadth of our review is that the overlap between the different explanatory factors proposed to explain contact-induced change is rarely explicit. Even when studies tackle similar topics (e.g. ‘foreigner-directed speech’, ‘grammar-based accommodation’ and ‘audience design’, cf. Uther *et al.* 2007; Fehér *et al.* 2019; Arnold *et al.* 2012, among many others), there seems to be little dialogue between studies looking into conceivably related phenomena. This finding may in part result from differences in research

traditions, and/or the existence of area-specific methods. In the above examples, Uther and colleagues observed interactions in lab settings, whereas Fehér and colleagues used an artificial language learning task. Generally speaking, researchers in second language learning tend to observe learners in the context of the classroom, whereas fieldworkers do not typically run perception experiments.

As for the nature of the data, some generalizations emerge from our survey. The first is that many studies are based on in-depth case studies of a single community in a contact situation (e.g. Türker 2000 for Turkish in Norway). Experimental studies, on the other hand, form the basis of the recent second language acquisition literature (e.g. Litcofsky *et al.* 2016), as well as of many bi-/multilingualism studies (e.g. Gyllstad & Wolter 2016 for collocational processing in second-language English speakers). A certain number of studies utilize census-type data, such as population size, number of speakers of a language, data on official languages, and so forth (e.g. Belew & Simpson 2018).

2.2 Types of explanatory factor

Using the procedure we outlined above enabled us to find a number of commonalities between all the different studies investigated nonetheless. We classified the various explanatory factors into four categories: *Cognitive Processes*, *Interactions between Individuals*, *Social Networks*, and *Macro-contexts of Language Use*. These are explained in detail below.

Explanatory factors pertaining to *Cognitive Processes* rely on domain-general processes, such as memory, categorization, perceptual saliency, etc. Proposals classified within this type of explanatory factor tend to focus on the individual as the agent of change, for instance by suggesting that individual learners' inability to hear a phonological distinction may lead to change in the L2 phoneme inventory. The data used in this type of explanatory factor generally come from experiments (e.g. Litcofsky *et al.* 2016), but also from language corpora, and from case studies (e.g. Blevins 2017). Phonological variables tend to predominate in proposals evoking Cognitive Processes. This is the case of Blevins (2017), where 'perceptual magnet' effects are evoked to explain areality in the distribution of sound patterns. Perceptual effects are typically used to account for how phonetic prototypes

function in perception (cf. Kuhl 1991; Kuhl *et al.* 2008). Many of these proposals also apply to other linguistic domains, as in the case of ‘metatypy’. Metatypy refers to syntactic and semantic changes that occur in bilingual communities, in which the less dominant language typically changes based on patterns in the more dominant one (Ross 2007: 116).

Explanatory factors classified as *Interactions between Individuals* also tend to focus on domain-general phenomena. One key difference between these factors and Cognitive Processes is that here the focus is on how individual speakers use language to interact with one another. That is, explanations in this category describe change as a product of interactions. For instance, in the case of ‘foreigner-directed speech’, native speakers would avoid using infrequent or complex constructions when talking to foreigners in an attempt to facilitate communication (see Rothermitch *et al.* 2019 and references therein). The data used in this type of proposal generally derive from case studies (e.g. Ferguson 1975; Berdicevskis 2020), although experiments also figure prominently (e.g. Uther *et al.* 2007; Weatherholtz *et al.* 2014; Chun *et al.* 2016). Studies on adaptive changes resulting from interactions between individuals commonly address morphosyntactic (Weatherholtz *et al.* 2014; Chun *et al.* 2016; Fehér *et al.* 2019) and phonological variables (Uther *et al.* 2007), even though proposals about accommodation in the lexicon have also been made (e.g. Ferguson 1975).

Proposals belonging within the *Social Networks* factor focus on the pathways of information flow between individuals and groups. The investigation of the diffusion of linguistic variables is prominent, and it emphasizes both within-group and between-group communication dynamics. The data used in these studies come from agent-based modeling work (e.g. Fagyal *et al.* 2010; Clem 2016), as well as fieldwork, and from case studies (e.g. Milroy & Milroy 1985; Lippi-Green 1989). The type of data discussed in Social Networks proposals vary widely, depending on the researcher’s focus. Social Network proposals seem readily amenable to the study of many types of variables, although replicability may be affected by the highly specific scenarios discussed.

Proposals belonging within *Macro-contexts of Language Use* make reference to the broadest contexts of language use. Such proposals tend to include socio-historical and socioeconomic variables, for instance

when discussing the role of colonization in the adoption of an official language (e.g. Spolsky & Lambert 2005), or the degree of language vitality (Mufwene 2017 and related response articles). Proposals in the Macro-contexts of Language Use often employ census-type data, meaning population size, official language, language of instruction, etc., although some sources also describe case studies. These studies tend to focus on post-hoc analyses, such as in the examination of the impact of ceremonial/religious language on various vernaculars (e.g. Fudge 2005). Because they are so general, Macro-contexts of Language Use explanations lend themselves to the study of most linguistic variables, with lexical variables more readily analyzable.

Uncovering, cataloguing, and grouping the myriad of explanatory factors scattered across related subfields constitutes a first attempt at proposing a framework of linguistic adaptation that better reflects the range of outcomes we encounter as researchers. At present, we are exploring possible overlaps between the four types of explanatory factors, with the ultimate goal of developing a comprehensive model of the mechanisms behind contact-induced change. We remain aware that this future model would still need testing. On the other hand, the information presented in this section has already proven a useful tool in the development of our sociolinguistic questionnaire (Section 5). In the next section we discuss further ways through which our study intends to build bridges across domains, namely by discussing comparative sociolinguistics.

3. Comparing sociolinguistic scenarios

In order to research whether linguistic structures adapt to sociolinguistic context across languages, we need to bridge the methodological approaches in language typology and sociolinguistics. A major challenge in this endeavor is the broadly differing methodological traditions in the two disciplines. Our solution is to develop an etic approach to language variation that is based on expert assessments.

Sociolinguistic research tends to focus on naturalistic language data. The data is annotated for language-specific grammatical features and for sociolinguistic features related to individual language users. Other data sources include interviews and questionnaires. The linguistic and social categories used in sociolinguistic research are largely

based on shared norms and have a social reality to the members of the community.

Typological research, on the other hand, focuses on linguistic diversity and largely uses reference grammars as data sources. Comparison across languages is typically based on tools defined by researchers that abstract away from language-particular categories (e.g., Stassen 1985; Haspelmath 2010). The tools used in comparison are thus created by the researcher. Although this approach is the state-of-the-art in language typology, issues about comparability are constantly being discussed (see, e.g. the discussion in special issue 20/2 in the journal *Linguistic Typology*). In typological research linguistic structures are generally analyzed into types that emerge from the variation. Yet, classifying linguistic structures into types is part of the typological method that relies on grammatical descriptions as data sources, rather than an end in itself.

Corpus data have not been as easily available in typology as in sociolinguistics, but the field has been changing recently towards the usage of naturalistic corpus data (e.g., Levshina 2019; Gerdes et al. 2021). This change has been boosted by the greater availability of multilingual parallel texts (e.g. Cysouw & Wälchli 2007) and annotated multilingual corpora, such as Universal Dependencies (Nivre *et al.* 2018). Because we aim at comparing languages across the world, a corpus-driven approach is not feasible for us, owing to the relatively poor availability of systematically annotated corpora especially for most minority languages in the world (but see e.g. the DoReCo project that is beginning to rectify this; Paschen et al. 2020).

In the current project, we develop an approach that bridges sociolinguistic and typological variation by using expert judgments as the basis for comparative data analyses. As mentioned above, the standard data source in typological research is reference grammars. These descriptive works are based on expert judgments about linguistic facts of the language in question. To match this approach in the sociolinguistic part of the project, we rely on expert judgements also for our sociolinguistic data. These data will be collected by using a sociolinguistic questionnaire that will be filled in by field experts collaborating with us. The sociolinguistic questionnaire is described in more detail in Section 5, but we elaborate on some key comparative principles behind it already here.

In our research design, the starting point is the mechanisms of language change, that is, the explanatory factors described in Section 2. We first analyze and classify different explanatory factors into groups and then start formulating broad questions about sociolinguistic factors related to those groups. For instance, related to cognitive processes the questionnaire asks to what extent children are exposed to multilingualism in language acquisition, and related to network structure we ask about the frequency of interaction between groups (e.g. Trudgill 2011).

However, broad questions yield broad answers. We assume that the social sphere is multidimensional and that this needs to be taken seriously in the research design to arrive at more informative answers. In our approach, we try to capture the multidimensionality of social contact by breaking down broad issues into more fine-grained ones. This process results in asking several questions, for instance, related to social networks and not just overall in the community but separately across six predefined social domains. These six social domains (local community, family & kin, social exchange & marriage, trade, labor, and knowledge) are described in Section 5.

We can illustrate the procedure by describing how we operationalize the distinction between ‘esoteric’ and ‘exoteric communication’, given that this distinction has been hypothesized to influence language use (Thurston 1987; Wray & Grace 2007; Givón 2009; Trudgill 2011). Esoteric communication takes place among intimates within a small group which means that the interlocutors tend to share much information as well as many norms. Exoteric communication, on the other hand, takes place mostly among strangers within a large group. In this type of communication, there tends to be less shared information between people. The hypothesis is that broad communication types may foster different patterns of interaction and linguistic use, thus ultimately impacting the evolution of different linguistic structures.

The distinction between esoteric and exoteric communication types is captured partly by the labels ‘dyad vs. group communication’ in our approach. First, we define these distinctions as universally applicable (the numbers given below are based on findings from experimental studies on communication; e.g. Fay *et al.* 2000; Fay & Ellison 2013):

- Dyadic communication refers to communication involving up to four people
- Group-based communication refers to communication involving ten or more people

These definitions are essentially etic and, in this sense, analogous to ‘comparative concepts’ in language typology (e.g. Haspelmath 2010).

We then formulate our questions so that they yield Likert-scale responses to probe variation a little more carefully than when asking mere categorical questions:

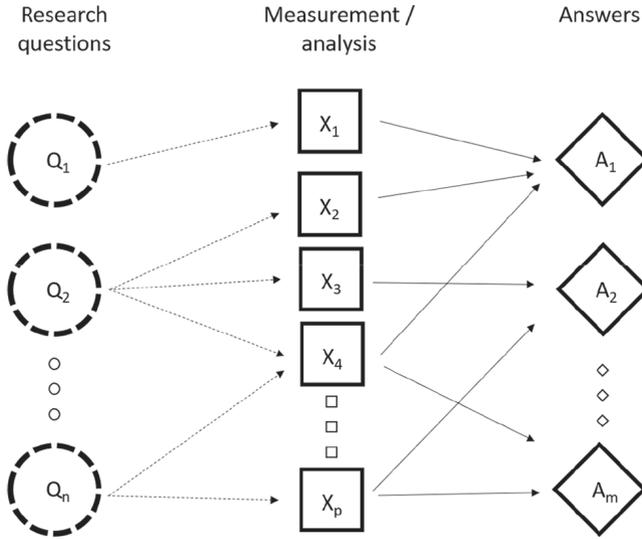
- Question: Are interactions between people in the community typically dyadic or group-based?
- Response options to the question
 - mostly dyadic (~highly esoteric)
 - somewhat dyadic
 - mixed
 - somewhat group-based,
 - mostly group-based (~highly exoteric)

However, this question-response frame by itself only addresses the type of communication in the whole group and thus potentially bypasses group-internal variation. In order to address group-internal variation, we repeat this question across the six predefined social domains. Other questions on social network structure asked in those domains are related to the following:

- Frequency of interaction (strength of connections in social networks)
- Amount of time spent in interaction (strength of connections)
- Effort in reaching other people in the network (strength of connections)
- Type of relationship between the interacting people: friends vs. enemies (nature and strength of connections in social networks)

The approach is schematically summarized in Figure 1.

Figure 1 - Schematic representation of the research process in typological approach to comparative sociolinguistic research (adapted from Vehkalahti 2019: 122)



To summarize, the process starts with designing the broad questions. These questions are then operationalized by breaking them down into more fine-grained questions which are used for data collection and analysis. Answers to the broad questions can be arrived at by aggregating the measurements in various ways. First, measurements can be aggregated by sets of questions which are related to the explanatory factors. Second, they can be aggregated by social domain. Third, they can be aggregated in other ways that may be theoretically well-motivated, or overall, to form a bird's-eye-view.

This approach allows us to take seriously the range of sociolinguistic contexts within a community, all the while approaching it from an etic perspective for comparative purposes, as is also done in state-of-the-art approaches to comparison in language typology. In the next section we describe how we sought to include a varied selection of sociolinguistic contexts through our sampling methodology.

4. *Sampling in sociolinguistic typology*

Investigating linguistic adaptation from a typological perspective while comparing sociolinguistic scenarios requires a diverse sample of speech communities from around the world. The sampling methodology developed within the GramAdapt project departs from existing sampling methods in language typology (for overviews, see Bakker 2011 and Miestamo *et al.* 2016.) in that we use sets of languages in contact, rather than individual languages, as the sampling unit. By definition, language contact involves two or more communities in interaction. Thus, understanding contact interactions and their repercussions on language structure requires that the unit of comparison include more than one language.

Our project uses established conventions in language contact research to define the internal structuring of our multi-language sampling units. The contact literature frames contact phenomena as pairwise interactions. In each of these pairs, one language counts as the potential recipient, and the other as the potential source of contact effects (Winford 2010: 171). We adopt this pairwise representation of contact scenarios and define the two languages in each pair as the *Focus Language* and *Neighbor Language*, respectively.

The *Focus Language* is the language whose potential linguistic adaptations we study. The *Neighbor Language* is a language identified in the reference materials to be in contact with the Focus Language. Specifically, the project investigates the Neighbor language's influence on the Focus Language by considering both patterns of language structures and social contact. Our sampling units also include a third language, termed the Benchmark Language, which allows us to disentangle areal diffusion from inheritance in the Focus Language. The Benchmark Language is genealogically related to the Focus Language, but is not in contact with either the Focus or the Neighbor Language. As such, the Benchmark Language serves as a parameter against which to test the impact of contact on the Focus language. Examples are given in Table 2.

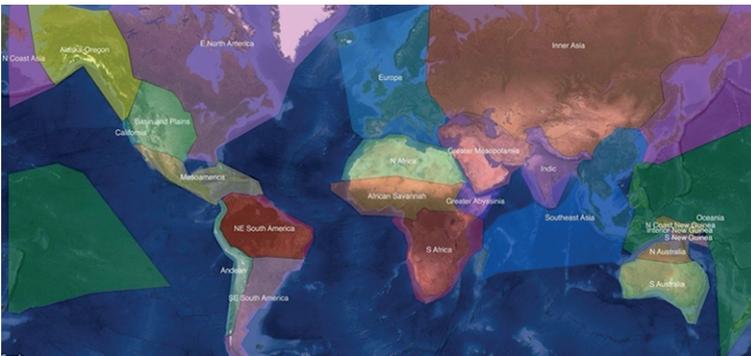
We are aware that the analysis of contact situations through the sets we propose has limitations. For instance, contact influence on the Focus Languages may come from other languages than those selected as Neighbor Languages in a given set. Moreover, the Neighbor Languages themselves may be affected by contact with the Focus Languages (and others). However, given that language contact minimally occurs between users of two languages, limiting our perspective to contact influences on

the Focus Language provides us with a unit of comparison that is schematic enough to be in principle applicable to any contact scenario worldwide. Similarly, restricting the choice of Benchmark to only one language per contact set only offers a partial view of the diachronic processes behind the retentions or innovations of linguistic features in the Focus Language. Nevertheless, sampling one Benchmark per contact set among the Focus Language's own relatives provides us with some measure of comparison that is external to the contact situation at stake, while at the same time controlling for genealogical relatedness.

To summarize, the proposed sampling technique establishes a principled way of capturing contact scenarios for the purpose of worldwide typological comparisons. We argue here that this is a crucial first step towards large scale studies of linguistic adaptation.²

Four sets of external criteria guided the compilation of the sample: (1) geographical area, (2) independently reported contact scenarios, (3) genealogical distance between languages in contact, and (4) availability of living experts to collaborate with. These four criteria allowed us to build a typological dataset that is geographically and genealogically stratified, in addition to being independent of our own assessment of a given contact situation. Each criterion is explained in more detail below.

Figure 2 - *The 24 Autotyp areas (Bickel et al. 2017, used under CC-BY 4.0 license)*



² In a way, all sampling methods in language typology involve some degree of coarseness in the framing of the units of comparison. For instance, probability samples, which are used to test statistical tendencies in the worldwide distribution of language structures, are typically constructed by extracting one (or a few) language(s) per genealogical unit. This tends to restrict the representation of the range of linguistic diversity attested in a given language family to the data point(s) chosen for that particular family.

The first criterion, geographical area, uses the 24 geographical areas established in the Autotyp database (Nichols *et al.* 2013; Bickel *et al.* 2017). The Autotyp areas (see Figure 2) derive from archeological, anthropological, historical, and genetic data, and are thus established independently of linguistic features (Nichols *et al.* 2013: 6). As such, using the Autotyp areas allows us to randomize language selection in a way that is blind to the goals of our own study. The sample currently consists of two sets of three languages for most Autotyp areas, totaling 150 languages. Two areas, Northeast South America, and Southern New Guinea, each provide three sets given the high degree of linguistic diversity found in South America and New Guinea as a whole (e.g. Dahl 2008; Hammarström 2016). The three language sets sampled for Southern New Guinea are shown in Table 2.

Table 2 - *Language sets sampled for the Trans Fly contact zone in Southern New Guinea (ISO 639-3 codes and language families are shown in parentheses)*

<i>Autotyp Area (Source)</i>		<i>Focus</i>	<i>Neighbor</i>	<i>Benchmark</i>
<i>Southern New Guinea</i> Evans <i>et al.</i> (2018)	<i>SET 1</i> ³	Nen (nqn; Yam)	Idi (Idi; Pahoturi River)	Yei (jei; Yam)
	<i>SET 2</i>	Coastal Marind (mrz; Anim)	Marori (mok; Marori)	Warkay-Bipim (bgv; Anim)
	<i>SET 3</i>	Koiari (kbk; Kolarian)	Motu (meu; Austronesian)	Ese (mcq; Kolarian)

The second criterion, namely independently reported contact scenarios, guides the choice of which languages to analyze from each area. This criterion reflects descriptions of language contact in the literature by experts in given areas of the world, such as in macro-area surveys of contact situations and/or areal linguistics. For instance, the surveys of languages of Southern New Guinea by Evans (2012) and Evans *et al.* (2018) describe several cases of contact situations in the area, from which we drew our own sets.

³ As with all other sampled sets, we use Glottolog as a reference system for the genealogical classification of the languages of this set. However, we remain aware that, in this particular case, Glottolog's internal groupings for the Yam family do not completely overlap with state-of-the-art comparative reconstructions (cf. Evans *et al.* 2018: 68)

The choice of the three languages also rests on the third criterion behind our sampling procedure: genealogical distance. Each set in our sample includes only Focus and Neighbor Languages that belong to distinct language families as classified in Glottolog (Hammarström *et al.* 2020). The choice of genealogically unrelated languages helps ensure that linguistic effects of the Neighbor Language on the Focus Language stem from the contact situation. On the other hand, the Benchmark Language is chosen among the Focus Language's relatives, with the added condition that it must not be part of the Focus/Neighbor contact zone.⁴ The degree of relatedness between Focus and Benchmark languages varies across the sampled sets, depending on the genealogy and contact history of individual language families. For instance, considering set 3 in Table 2, according to Glottolog's classification, Koiari and Ese, the Focus and Benchmark languages, belong to two distinct upper level subdivisions of the Koiaran family. Koiari is part of the Koiaric subgrouping while Ese is part of the Baraic subgrouping and is spoken outside the contact zone where the Focus and Neighbor language are found.⁵

The fourth and final selection criterion is the availability of experts who could describe the contact scenario between Focus and Neighbor Language. This criterion follows from the fact that collaboration with experts is at the core of the data collection through our sociolinguistic questionnaire (Section 5). When selecting our pool of experts, we also strived to involve researchers from the communities under study, in an attempt to engage with different academic cultures and more local perspectives on the contact scenarios at stake.

Following the procedure outlined above, we identified 50 contact sets from all parts of the world and across all Autotyp Areas.⁶ The sample currently includes 50 Focus Languages distributed across 36 lan-

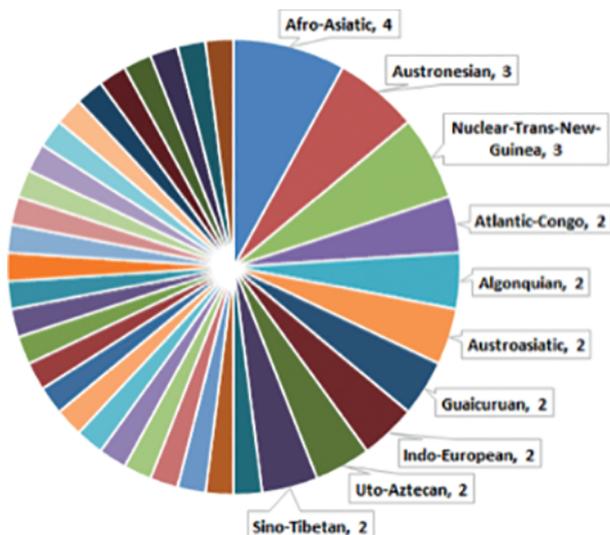
⁴ One consequence of this setup is that while Neighbor Languages can be isolates (e.g. Basque in the Gascon-Basque contact pair), Focus Languages cannot. By establishing that Focus Languages should be compared with Benchmark Languages, isolates are by definition excluded.

⁵ We will apply the Diversity Value Method (Rijkhoff *et al.* 1993; Rijkhoff & Bakker 1998) on all pairs of Focus and Benchmark languages as a measure of genealogical relatedness. This relatedness score will be used as one of the independent variables to factor into the analyses of the linguistic and sociolinguistic data.

⁶ Sample size was set at 50 sets (i.e. 150 languages in total) so as to maximize diversity in sample composition, while keeping a manageable workload given the time frame of the project.

guage families (as per Glottolog's classification), with an average of 1.4 languages per family.⁷ While languages from large and geographically widespread language families (e.g. Afro-Asiatic, Austronesian, Nuclear-Trans-New-Guinea) are represented more than once in the sample in different roles (i.e. as Focus, Neighbor, or Benchmark languages), the majority of the sampled language families are small. Therefore, most language families are represented only once in our dataset. Figure 4 illustrates our sample by reporting the distribution of Focus Languages per language family. The figure also details the number of Focus Languages selected for the ten language families with more than one representative included in the sample.

Figure 3 - *Distribution of Focus Languages per language family, detailing the number of Focus languages from families with more than just one representative in our sample. Focus Languages Genealogical Affiliation (N=50)*



The selection process was fully independent of the linguistic variables that we use to assess the linguistic outcomes of a given contact situation (Section 6). The language selection process was also independent of the

⁷ We only report on the Focus Languages here because the choice of Neighbor Languages is still under discussion for a few sets. As mentioned above, Benchmark languages belong to the same families as Focus languages.

social history of a given contact situation, which is what we aim to study in depth through the sociolinguistic questionnaire, explained in the next Section. To summarize, the sampling technique presented here is, to the best of our knowledge, the first attempt to develop a typological dataset that is specifically designed for worldwide studies of linguistic adaptation. As we hope to have shown, the methodology is transparent and replicable, and has the potential to provide a tool for investigating the influence of sociolinguistic environments on the distribution of linguistic diversity.

5. Sociolinguistic questionnaire

As stated in Section 3, this project attempts to bridge sociolinguistic and typological variation by using expert judgments as the basis for comparative data analyses. The GramAdapt questionnaire is a practical manifestation of this attempt.

The questionnaire comprises two parts. The first part is the Overview Questionnaire, which is designed to match prior studies investigating correlations between linguistic and macro socio-cultural factors. These macro-factors include speaker population (Lupyan & Dale 2010), degrees of political complexity (Currie & Mace 2009), and types of marriage systems (Bowern 2010 in hunter-gatherer societies). The second part is the Domains Questionnaire, which will be the focus for the remainder of this paper.

The Domains Questionnaire design is based on the well-established finding that different kinds of interactional situations beget different kinds of linguistic behaviors; known in the literature as “the domains of language use” (Fishman 1965). We have identified six social domains that typically have norms and modes of interaction particular to that domain, and are often attested as domains of social contact.

The six social domains are: local community, family & kin, social exchange & marriage, trade, labor, and knowledge. We developed operational definitions for each domain with the aim to make them as cross-culturally applicable as possible, which will also aid in the development of comparative tools as outlined in Section 3. Given the range of societal diversity we see across contact situations in space and time, the domains must be flexible enough to capture the dynamics of small and egalitarian communities, right through to communities that are part of larger, socially stratified and administratively complex societies.

While there are inevitable overlaps between the domains when considering real communities, we have attempted to create an operational definition for each domain, which is as outlined below:

- **Local Community:** Concerning interactions beyond kin, and outside institutionalized practices related to trade, labor, and knowledge. For example, a village, a band, or a neighborhood in a town.
- **Family & Kin:** Concerning the interaction between members of family and kin. This domain typically encompasses child bearing and rearing, as well as food production and consumption.
- **Social Exchange & Marriage:** Concerning practices of exchange which regulate relationships between individuals and groups, within and across societies. This domain encompasses practices of gift and ceremonial exchange, as well as marriage exchange.
- **Trade:** Concerning transactions of objects and services. The prototype is a transaction of commodities. The mode of transaction can be monetary, barter-based, etc.
- **Labor:** Concerning practices and relationships that revolve around economic activity and production.
- **Knowledge:** Concerning knowledge transfer that is structured in culturally specific ways. This domain prototypically covers practices that revolve around education and religion.

As stated earlier, the goal of the Domains Questionnaire is to seek data on contact dynamics in each of the six domains along the four explanatory factors for contact-induced change illustrated in Section 2, that is cognitive processes, interaction between individuals, social networks, and macro-contexts of language use (see Section 2 and 3 for illustrations of each of these factors and their relation to our comparative sociolinguistic approach). The data will be collected for those domains where social contact occurs between Focus and Neighbor groups. The response across the social domains would then scale up to a general profile of social contact, from the bottom-up.

To give an illustration, for the explanatory factor Social Networks, we have seven questions designed to get a sense of social network scope. For example, one question asks about the frequency of interaction between Focus and Neighbor Group peoples (question S1 in table 4). Another asks about interaction type: whether it

is dyadic, or group-based and broadcast-like (Section 3). For every domain where social contact occurs, the respondent will provide an answer to these questions.

If, for a particular subset of the Domains Questionnaire there is no contact between the groups, the respondent will simply skip that particular domain. Once the questionnaire is filled, we will have an aggregate view of how dense contact is across domains. We may surmise that the more social domains where Focus and Neighbor Group peoples interact, the denser the contact.

We illustrate how the Questionnaire is structured with examples of two situations: one in New Guinea and the other in India. In one situation, we have social contact between speakers of Nen (Yam), and Idi (Pahoturi River) in southern Papua New Guinea. In the other situation, we have social contact between speakers of Marathi (Indo-Aryan) and Kannada (Dravidian) in Kupwar, India. The socio-cultural configuration of the two situations is distinct. In the New Guinea case, the groups are non-hierarchical, and based on subsistence horticulture. In India, we have a caste-based stratified society which is part of a larger socio-political complex. In New Guinea, the total number of speakers is quite small: 350 for Nen, and 800 for Idi (Evans *et al.* 2018: 645-646). The total number of Marathi and Kannada speakers in Maharashtra state is 77.5 million and 1 million respectively (Office of the Registrar General & Census Commissioners 2011). More details on the socio-cultural configurations of the two situations is presented in Table 3.

Table 3 - *Overview of some societal-demographic characteristics of Nen/Idi contact pair, and Marathi/Kannada contact pair*

<i>Contact Pair</i>	Focus Group = Nen (Yam); Neighbor Group = Idi (Pahoturi River)	Focus Group = Marathi (Indo-Aryan); Neighbor Group= Kannada (Dravidian)
<i>Autotyp Area</i>	Southern New Guinea	Indic
<i>Language and social affiliation</i>	Äkamar tribe people claim Nen as their language. Gunduma people claim Idi as their language	Low-caste Hindus speak Marathi, land owning Jains and Lingayat crafts people speak Kannada

<i>Socioeconomic hierarchies and subsistence pattern</i>	Non-hierarchical, subsistence horticulture	Caste-based professional hierarchy, complex agriculture
<i>Relationship to larger socio-political structures</i>	Historically not part of a centralized state until effective colonization in the 1960s, currently national infrastructure is mostly absent	Historically parts of various empires, currently part of Maharashtra state of the Republic of India with Marathi the official language of education and administration in Maharashtra state
<i>Language ideologies</i>	Egalitarian multilingualism (as defined by François 2012) with a relative absence of major languages (i.e. Tok Pisin)	Caste and religion-based linguistic divisions
<i>Speaker population</i>	Approximate total speaker population: Nen = 350 Idi = 800	Approximate total speaker population: Marathi = 77.5 million in Maharashtra state; 83 million in the whole of India Kannada = 1 million in Maharashtra state; 43.5 million in the whole of India

Let us consider three social domains: Local Community, Family & Kin, and Trade. Under the operational definitions we produced for the questionnaire, the New Guinea situation shows social contact in two domains: Local Community, and Family & Kin. In our definition of trade, we emphasize the fact that a given transaction presupposes an expectation of immediate or future return. This configuration does not apply to the New Guinea scenario.⁸ The respondent will therefore answer the questions for Local Community and Family & Kin, but not for Trade. The India situation, on the other hand, shows social contact in Local Community and Trade, but not Family & Kin. Language is a property of caste membership (Gumperz & Wilson 1971; Kulkarni-Joshi 2016), and people of different castes do not intermarry. The respondent will therefore answer the questionnaire for the domains of local community and trade, but not family & kin.

Table 4 provides an example of responses for the Nen/Idi and Marathi/Kupwar contact situations in the domain of Local

⁸ The social domain concerned with what we call “social exchange” would cover situations like the New Guinea case where relationship building is either an explicit or crucial motivator of conducting exchange.

Community. The possible responses for the Social Network set of questions are based on a five-point scale,⁹ as shown earlier on in Section 3. Eri Kashima provided the responses for the purpose of this demonstration. She is an expert of Nmba (ncm; Yam) and has worked with Nen speakers in the context of studying the sociolinguistics of the area. Eri Kashima also filled out the Marathi/Kannada response for the purposes of this demonstration, based on Gumperz & Wilson (1971) and Kulkarni-Joshi (2015; 2016). The final version of the questionnaire will be filled out by an expert who has worked in the Marathi/Kannada context. The sample answers for Local Community indicate that the Nen/Idi contact situation may be denser than Marathi/Kannada. These responses will form one portion of our bottom-up characterization of social contact in these situations, which can then be compared across the sample set.

One major challenge in designing the questionnaire has been setting the time-frame of the questionnaire response. In order to test for linguistic adaptation, one would need to know about the contact situation before present; since linguistic changes visible at the present would have adapted to a given socio-historical context in the past. Given the challenges in generalizing across the time-depth of contact scenarios worldwide, we refrain from establishing any a priori chronological cut-off points for the contact situations in this questionnaire. Instead, we ask respondents to (a) assess the duration of contact between Focus and Neighbor Group in a given social domain, and (b) identify the time frame of densest contact between Focus and Neighbor Group in said social domain. The questions should be answered from the perspective of this time frame. In doing so, we hope to gain an understanding of the contact situations at stake that is maximally entrenched in the specifics of their linguistic and social ecologies. There are multiple ways in which this time-frame issue could be dealt with, but our solution strikes a balance between the availability of data and the diversity of contact situations from a global perspective.

⁹ Not all of the questions, however, are based on a five-point scale. There are many questions with binary “yes/no” answers.

Table 4 - *An example of answers to the Social Network questions (S-Set). These sets of questions require an answer on a five-point scale. The ID number of questions (S1, S3 etc.) is non-contiguous as question S2 was retired during the design process. The expression of the question in this demonstration shows the essence of the question, rather than the final formulation in the questionnaire*

<i>Network Structure Question</i>	<i>Nen & Idi (New Guinea)</i>	<i>Marathi & Kannada (India)</i>
<i>S1: How often do Focus Group people typically interact with Neighbor Group people?</i>	3	4
<i>S3: How many people are typically involved in interactions between Focus Group and Neighbor Group? Is it more-or-less dyadic or group-based?</i>	4	2
<i>S4: How physically proximate to each other are Focus Group people and Neighbor Group people in this domain?</i>	4	4
<i>S5: How friendly are Focus Group people and Neighbor Group people in this domain?</i>	5	3
<i>S6: What is the proportion of total Focus Group people who have opportunities for contact with Neighbor Group people</i>	5	2
<i>S7: What is the proportion of total Neighbor Group people who have opportunities for contact with Focus Group people</i>	3	2
<i>Network Density Score for Local Community</i>	24	17

6. *Linguistic variables*

Aside from collecting data on social contact between Focus and Neighbor Groups as detailed in Section 5, we will also work with a selection of linguistic variables from different domains of language structure ranging from phonology, morphosyntax, and the lexicon. The coding for each of these variables will be based on established literature on relevant patterns of crosslinguistic variation and diachronic change in these domains. The variables will be used to test hypotheses about linguistic adaptation in the languages of the sample; that is, to investigate whether any structural features of the Focus Languages may be attributed to contact with the respective Neighbor languages or whether other processes (i.e. plain retention and/or language internal evolution) may be at stake. The linguistic variables are chosen among renown cross-linguistically common structures in the domain

of phonology, morphosyntax, and the lexicon. While the variable selection process is still ongoing, we have a preliminary plan of action for the domains that we would like to consider.

Within phonology, we will be looking at suprasegmental patterns, such as syllable structure and word stress patterns. As recently demonstrated in Napoleão de Souza & Sinnemäki (revised), suprasegmental features represent a promising testing ground to investigate phonological change in contact situations. Within morphosyntax, we will be focusing on nominal number and locus of marking. Number is the most frequent nominal category cross-linguistically (Corbett 2000) and is also relatively well studied from a typological and language contact perspective, which facilitates both variable selection and data collection using reference materials (for a discussion of number marking and language contact dynamics see, for instance, Roberts & Bresnan 2008 and Igartua 2015). The same could be said about locus of marking (Nichols 1992). Locus of marking concerns syntactic relations and how they are morphologically marked within phrases and the clause. Syntactic relations are often morphologically marked in languages either on the head or the dependent of the construction, and these markings are prone to change especially under heavy language contact (Roberts & Bresnan 2008). Finally, in the domain of the lexicon, we will study demonstrative systems. Our focus will be on adnominal demonstratives. These are often mentioned in the contact linguistics literature, but have never been systematically studied from the perspective of language adaptation, which we will be testing for the first time.

Working on a selection of linguistic features from a diverse range of structural domains increases the chance that at least some of the variables of choice will be meaningful for the purpose of identifying and describing processes of adaptation. Moreover, this procedure allows us to test whether the likelihood of adaptation differs across domains – e.g. if prosodic features turn out to be more likely to undergo adaptive changes than demonstrative systems or number systems. Finally, the method enables us to investigate whether the types of attested language changes, as well as their frequency of occurrence, vary depending on the length and intensity of contact, which we estimate through the sociolinguistic questionnaire.

All sampled languages will be coded for the same set of linguistic features. This procedure allows us to run comparable analyses across sampled languages and contact sets. Through these comparisons, we can then assess the possibility that some variables may be more relevant to certain contact sets than others. We are aware that linguistic adaptations may occur elsewhere than in those domains we have chosen to investigate.

A crucial issue in the linguistic data collection process concerns variable design. Much of the research on linguistic adaptation has focused on number of distinctions, for instance, in phoneme inventory sizes (Hay & Bauer 2007), number of cases (Bentz & Winter 2013; Sinnemäki 2020), and number of gender distinctions (Sinnemäki & Di Garbo 2018; Dahl 2019). However, as suggested in recent research (cf. Sinnemäki & Di Garbo 2018; Verkerk & Di Garbo *accepted*, 2021, regarding grammatical gender), counting the number of distinctions in a grammatical domain may not always be the best way to assess patterns of linguistic adaptation.

An alternative to the number-of-distinctions approach is to consider the processes of restructuring that language structures undergo from a cross-linguistic perspective, and how they may correlate with different types of sociolinguistic scenarios. We embrace this alternative approach, which has also been validated in recent research on linguistic adaptation and contact-induced change.

With respect to linguistic adaptation in morphosyntax, researchers have shown that contact situations characterized by high proportions of adult L2 learning favor processes of restructuring that increase the transparency and compositionality of morphosyntactic paradigms (Kusters 2003; Trudgill 2011; Kempe & Brooks 2018). For instance, patterns of grammatical gender marking only marginally shaped by semantic criteria may become fundamentally restructured around the encoding of animacy distinctions. By studying the evolution of gender systems in northwestern Bantu languages, Di Garbo & Verkerk (*accepted*, 2021) and Verkerk & Di Garbo (*accepted*, 2021) find that these highly transparent gender systems abound in languages with a history of intense language contact and/or language shift¹⁰.

¹⁰ Previous studies that looked at the sociolinguistic typology of grammatical gender by using the number of gender distinctions as the linguistic variable of interest were not able to show any such effect (Sinnemäki & Di Garbo 2018; Dahl 2019).

Similarly, Napoleão de Souza & Sinnemäki (revised) demonstrate that looking at processes of restructuring in suprasegmental features may be more informative than a simple assessment of the presence vs. absence of some phonological variables. The authors claim that focusing on processes may advance our understanding of the impact of language contact on phonological structure.

In view of the evidence presented above, for each of the linguistic variables of choice, we will develop our coding design in a way that is informed by research in typology, historical linguistics, sociolinguistics, and studies of bi-/multilingual language use.

7. Concluding remarks

In this paper we outlined the approach developed within the GramAdapt project with the aim of establishing appropriate concepts and methods for investigating the relationship between languages and their sociolinguistic environments. The proposed framework can be used to study sociolinguistic correlates of linguistic diversity and language change in three ways: (1) through the analysis of causal factors related to language change, (2) through a novel sampling technique simultaneously addressing selection of communities, sociolinguistic features, and linguistic features, and (3) through generalizations from empirically-grounded cross-cultural and cross-linguistic comparisons. It is our hope that this approach to worldwide comparisons of language structures and communities will set a new ground for the typological study of linguistic adaptation.

Acknowledgements

This research has received funding by the European Research Council (ERC), grant no 805371 to Kaius Sinnemäki (PI). We wish to acknowledge the audience of the SLI workshop on “Sociolinguistics and linguistic typology: towards an integrated approach to the study of linguistic variation” (September 2020) and, in particular, the workshop organizers and editors of this volume Silvia Ballarè and Guglielmo Inglese. For feedback on the research design, we are especially grateful to Maria Khachatryan, Olesya Khanina, Friederike Lüpke, Susanne Michaelis, Brigitte Pakendorf, Ksenia Shagal, and

Max Wahlström, as well as to the audience of the GramAdapt online seminar series (November – December 2020). For help and support with the manuscript layout, we thank Janne Loisa. The usual disclaimers apply.

Bibliography

- Atkinson, Quentin. 2011. Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science* 332(6027). 346-349.
- Arnold, Jennifer E., M. Kahn, Jason & C. Pancani, Giulia. 2012. Audience design affects acoustic reduction via production facilitation. *Psychonomic bulletin & review* 19(3). 505-512.
- Bakker, Dik. 2011. Language Sampling. In Song, Jae Jung (ed.) *Handbook of Linguistic Typology*, 100-127. Oxford: Oxford University Press.
- Belew, Anna & Simpson, Sean. 2018. The status of the world's endangered languages. In Rehg, Kenneth L. & Campbell, Lyle (eds.) *Oxford handbook of endangered languages*, 21-47. Oxford: Oxford University Press.
- Bentz, Christian & Winter, Bodo. 2013. Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change* 3. 1-27.
- Berdicevskis, Aleksandrs. 2020. Foreigner-directed speech is simpler than native-directed: Evidence from social media. *Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science*. 163-172.
- Bickel, Balthasar, Nichols, Johanna, Zakharko, Taras, Witzlack-Makarevich, Alena, Hildebrandt, Kristine, Rießler, Michael, Bierkandt, Lennart, Zúñiga, Fernando & Lowe, John B. 2017. *The AUTOTYP typological databases*. Version 0.1.0 <https://github.com/autotyp/autotyp-data/tree/0.1.0>
- Blasi, Damián E., Moran, Steven, Moisik, Scott R., Widmer, Paul, Dediu, Dan & Bickel, Balthasar. 2019. Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science* 363(6432). doi: 10.1126/science.aav3218
- Blevins, Juliette. 2017. Areal Sound Patterns: From Perceptual Magnets to Stone Soup. In Hickey, Raymond (ed.), *The Cambridge Handbook of Areal Linguistics*. 88-121. Cambridge: Cambridge University Press.

- Bowern, Claire. 2010. Correlates of Language Change in Hunter-Gatherer and Other 'Small' Languages. *Language and Linguistics Compass* 4(8). 665-679.
- Chun, Eunjin, Barrow, Julia, & Kaan, Edith. 2016. Native English speakers' structural alignment mediated by foreign-accented speech. *Linguistics Vanguard* 2(s1). 1-10.
- Clem, Emily. 2016. Social network structure, accommodation, and language change. *UC Berkeley Phonetics and Phonology Lab Annual Report*, 83-102.
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.
- Currie, Thomas E. & Mace, Ruth. 2009. Political complexity predicts the spread of ethnolinguistic groups. *Proceedings of the National Academy of Sciences* 106(18). 7339-7344.
- Cysouw, Michael & Wälchli, Bernhard. 2007. Parallel texts: using translational equivalents in linguistic typology. *STUF – Language Typology and Universals* 60(2). 95-99. <https://doi.org/10.1524/stuf.2007.60.2.95>.
- Dahl, Östen. 2004. *The growth and maintenance of complexity*. Amsterdam: John Benjamins.
- Dahl, Östen. 2008. An exercise in a posteriori language sampling. *STUF – Language Typology and Universals* 61(3). 208-220.
- Dahl, Östen. 2019. Gender: exoteric or esoteric? In Di Garbo, Francesca, Olsson, Bruno & Wälchli, Bernhard (eds.), *Grammatical gender and linguistic complexity, vol. I: General issues and specific studies*, 53-61. Berlin: Language Science Press.
- Di Garbo, Francesca & Verkerk, Annemarie. Accepted, 2021. *A typology of northwestern Bantu gender systems*. To appear in *Linguistics*.
- Evans, Nicholas. 2012. Even more diverse than we thought: the multiplicity of Trans-Fly languages. In Evans, Nicholas and Klamer, Marian (eds.) *Melanesian Languages on the Edge of Asia: Challenges for the 21st Century*, 109-149. Honolulu: University of Hawai'i Press.
- Evans, Nicholas, Arka, I. Wayan, Carroll, Matthew J., Choi, Yun Jung, Döhler, Christian, Gast, Volker, Kashima, Eri, Mittag, Emile, Olsson, Bruno, Quinn, Kyla, Schokkin, Dineke, Tama, Phillip, van Tongeren, Charlotte, & Siegel, Jeff. 2018. The Languages of Southern New Guinea. In Palmer, Bill (ed.), *The Languages and Linguistics of New Guinea: A Comprehensive Guide*, 640-774. Berlin: De Gruyter Mouton.

- Everett, Caleb, Blasi, Damián E. & Roberts, Seán G. 2015. Climate, vocal folds, and tonal languages: connecting the physiological and geographic dots. *PNAS* 112. 1322-1327.
- Fagyal, Zsuzanna., Swarup, Samarth, Escobar, Anna María, Gasser, Les, & Lakkaraju, Kiran. 2010. Centers, Peripheries, and Popularity: The Emergence of Norms in Simulated Networks of Linguistic Influence. *University of Pennsylvania Working Papers in Linguistics* 15(2). 81-90.
- Fay, Nicholas, & Ellison, T. Mark. 2013. The Cultural Evolution of Human Communication Systems in Different Sized Populations: Usability Trumps Learnability. *PLoS One* 8(8). <https://doi.org/10.1371/journal.pone.0071781>
- Fay, Nicholas, Garrod, Simon & Carletta, Jean. 2000. Group discussion as interactive dialogue of group monologue: The Influence of Group Size. *Psychological Science* 11(6). 481-486.
- Fehér, Olga, Ritt, Nikolaus & Smith, Kenny. 2019. Asymmetric accommodation during interaction leads to the regularisation of linguistic variants. *Journal of Memory and Language* 109. 104036.
- Ferguson, Charles A. 1975. Toward a characterization of English foreigner talk. *Anthropological Linguistics* 17. 1-14.
- Fishman, Joshua A. 1965. Who Speaks What Language to Whom and When? *La Linguistique* 1(2). 67-88.
- Fudge, Erik. 2005. Religion and language. In Keith Brown (ed.) *Encyclopedia of language and linguistics*. Elsevier.
- Gerdes, Kim, Kahane, Sylvain & Chen, Xinying. 2021. Typometrics: From implicational to quantitative universals in word order typology. *Glossa: A Journal of General Linguistics* 6(1). 17. <http://doi.org/10.5334/gjgl.764>
- Givón, Talmy 2009. *Syntactic Complexity: Diachrony, Acquisition, Neurology, Evolution*. Amsterdam: John Benjamins.
- Gumperz, John, & Wilson, Robert. 1971. Convergence and creolization: a case from the Indo-Aryan/Dravidian border in India. In Hymes, Dell H. (ed.), *Pidginization and Creolization of Languages*, 153-169. Cambridge: Cambridge University Press.
- Gyllstad, Henrik & Wolter, Brent. 2016. Collocational processing in light of the phraseological continuum model: Does semantic transparency matter? *Language Learning* 66(2). 296-323.
- Hammarström, Harald. 2016. Linguistic diversity and language evolution. *Journal of Language Evolution* 1(1). 19-29. <https://doi.org/10.1093/jole/lzw002>

- Hammarström, Harald, Forkel, Robert, Haspelmath, Martin & Bank, Sebastian. 2020. *Glottolog database 4.3*. Jena: Max Planck Institute for the Science of Human History. doi:10.5281/zenodo.3754591.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663-687. <http://www.jstor.org/stable/40961695>.
- Hay, Jennifer & Bauer, Laurie. 2007. Phoneme inventory size and population size. *Language* 83. 388-400.
- Hruschka, Daniel J., Christiansen, Morten H., Blythe, Richard A., Croft, William, Heggarty, Paul, Mufwene, Salikoko S., Pierrehumbert, Janet B., & Poplack, Shana. 2009. Building social cognitive models of language change. *Trends in Cognitive Sciences*, 13(11). 464-469. <https://doi.org/10.1016/j.tics.2009.08.008>
- Igartua, Ivan. 2015. From cumulative to separative exponence: Reversing the morphological cycle. *Language* 91(3). 676-722.
- Kempe, Vera & Brooks, Patricia J. 2018. Linking adult second language learning and diachronic change: A cautionary note. *Frontiers in Psychology* 9. doi:10.3389/fpsyg.2018.00480. <https://www.frontiersin.org/article/10.3389/fpsyg.2018.00480>.
- Kousters, Wouter. 2003. *Linguistic Complexity: The Influence of Social Change on Verbal Inflections*. Utrecht: LOT Dissertation Series.
- Kuhl, Patricia. 1991. Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories. Monkeys do not. *Perception & psychophysics* 50. 93-107. 10.3758/BF03212211.
- Kuhl, Patricia K., Conboy, Barbara T., Coffey-Corina, Sharon, Padden, Denise, Rivera-Gaxiola, Maritza, & Nelson, Tobey. 2008. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences* 363(1493). 979-1000.
- Kulkarni-Joshi, Sonal. 2015. Religion and language variation in a convergence area: The view from the border town of Kupwar post-linguistic reorganisation of Indian states. *Language and Communication* 42. 75-85.
- Kulkarni-Joshi, Sonal. 2016. Forty years of language contact and change in Kupwar: A critical assessment of the intertranslatability model. *Journal of South Asian Languages and Linguistics* 3(2). 147-174.
- Levshina, Natalia. 2019. Token-based typology and word order entropy: A study based on Universal Dependencies. *Linguistic Typology* 23(3). 533-572.

- Lippi-Green, Rosina L. 1989. Social network integration and language change in progress in a rural alpine village. *Language in Society* 18(2). 213-234.
- Litcofsky, Kaitlyn A., Tanner, Darren & van Hell, Janet G. 2016. Effects of language experience, use, and cognitive functioning on bilingual word production and comprehension. *International Journal of Bilingualism* 20(6). 666-683.
- Lupyan, Gary & Dale, Rick. 2010. Language structure is partly determined by social structure. *PLOS One* 5(1). 1-10.
- Lupyan, Gary & Dale, Rick. 2016. Why Are There Different Languages? The Role of Adaptation in Linguistic Diversity. *Trends in Cognitive Sciences* 20(1). 649-660.
- Miestamo, Matti, Bakker, Dik & Arppe, Antti. 2016. Sampling for variety. *Linguistic Typology* 20(2). doi:10.1515/lingty-2016-0006.
- Milroy, James & Milroy, Lesley. 1985. Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21(2). 339-384.
- Mufwene, Salikoko S. 2017. Language vitality: The weak theoretical underpinning of what can be an exciting research area. *Language* 93(4). 202– 223.
- Napoleão de Souza, Ricardo & Sinnemäki, Kaius. Revised. *Beyond segment inventories: Phonological complexity measures and suprasegmental variables in contact situations*.
- Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*. Chicago, IL: University of Chicago Press.
- Nichols, Johanna, Witzlack-Makarevich, Alena & Bickel, Balthasar. 2013. *The AUTOTYP genealogy and geography database: 2013 release*. <http://www.spw.uzh.ch/autotyp/>.
- Nivre, Joakim, Abrams, Mitchell, Agić, Željko, et al. 2018. *Universal Dependencies 2.3. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University*. <http://hdl.handle.net/11234/1-2895>.
- Office of the Registrar General & Census Commissioner. 2011. *Part A: Distribution of the 22 scheduled languages-India/States/Union Territories*. Retrieved from [https://censusindia.gov.in/2011Census/Language_MTs.html]
- Paschen, Ludger, Delafontaine, François, Draxler, Christoph, Fuchs, Susanne, Stave, Matthew & Seifart, Frank. 2020. Building a time-aligned cross-linguistic reference corpus from language documentation data (DoReCo). *Proceedings of the 12th Conference on Language Resources and*

- Evaluation (LREC 2020)*, 2657-2666. Marseille: European Language Resources Association. <http://www.lrec-conf.org/proceedings/lrec2020/pdf/2020.lrec-1.324.pdf>
- Rijkhoff, Jan & Bakker, Dik. 1998. Language sampling. *Linguistic Typology* 2-3(3). 263-314.
- Rijkhoff, Jan, Bakker, Dik, Hengeveld, Kees & Kahrel, Peter. 1993. A method of language sampling. *Studies in Language* 17. 169-203.
- Roberts, Sarah and Joan Bresnan. 2008. Retained inflectional morphology in pidgins: A typological study. *Linguistic Typology* 12. 269-302.
- Ross, Malcom. 2007. Calquing and metatypy. *Journal of Language Contact* 1(1). 116-143.
- Rothermich, Kathrin, Harris, Havan L., Sewell, Kerry & Bobb, Susan C. 2019. Listener impressions of foreigner-directed speech: A systematic review. *Speech Communication* 112. 22-29.
- Sinnemäki, Kaius. 2020. Linguistic system and sociolinguistic environment as competing factors in linguistic variation: A typological approach. *Journal of Historical Sociolinguistics* 6(2). <https://doi.org/10.1515/jhsl-2019-1010>.
- Sinnemäki, Kaius & Di Garbo, Francesca. 2018. Language structures may adapt to the sociolinguistic environment, but it matters what and how you count: A typological study of verbal and nominal complexity. *Frontiers in Psychology* 9. doi:10.3389/fpsyg.2018.01141. <https://www.frontiersin.org/article/10.3389/fpsyg.2018.01141>.
- Spolsky, Bernard & Richard D. Lambert. 2005. Language planning and policy: Models. In Keith Brown (ed.) *Encyclopedia of language and linguistics*. Elsevier.
- Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell.
- Thurston, William R. 1987. Processes of Change in the Languages of North-Western New Britain. *Pacific Linguistics*. B-99. Canberra: Australian National University.
- Trudgill, Peter. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.
- Türker, Emel. 2000. *Turkish-Norwegian codeswitching: Evidence from intermediate and second generation Turkish immigrants in Norway*. Oslo: University of Oslo. (PhD Dissertation.)

- Uther, Maria, Knoll, Monja A., & Burnham, Denis. 2007. Do you speak E-NG-LI-SH? A comparison of foreigner-and infant-directed speech. *Speech Communication* 49(1). 2-7.
- Vehkalahti, Kimmo. 2019. Kyselytutkimuksen menetelmät ja mittarit [Methods and measures in questionnaire surveys]. Helsinki: University of Helsinki. <https://doi.org/10.31885/9789515149817>.
- Verkerk, Annemarie & Di Garbo, Francesca. Accepted, 2021. *Socio-geographic correlates of typological variation in northwestern Bantu gender systems*. To appear in *Language Dynamics and Change*.
- Weatherholtz, Kodi, Campbell-Kibler, Kahryn & Jaeger, T. Florian 2014. Socially-mediated syntactic alignment. *Language Variation and Change* 26(3). 387-420.
- Winford, Donald. 2010. Contact and borrowing. In Hickey, Raymond (ed.) *The Handbook of Language Contact*. 170-187. Somerset: John Wiley & Sons, Incorporated.
- Wray, Alison & Grace, George W. 2007. The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua* 117(3). 543-578. <https://doi.org/10.1016/j.lingua.2005.05.005>.

FABIO GASPARINI

Nominalizzazione indipendente in sudarabico moderno

Questo intervento vuole fornire una riflessione sui processi di insubordinazione in sudarabico moderno (Semitico, Afroasiatico). In particolare, oggetto di indagine saranno le cosiddette costruzioni ‘pseudorelative’ (Pennacchietti 2007) in Mehri e Baḥari, due varietà strettamente collegate geneticamente e sociolinguisticamente. La questione sarà riveduta applicando come modello di analisi il concetto di nominalizzazione indipendente (‘stand-alone nominalization’), termine con cui si indica un processo di nominalizzazione a livello frasale in contesto sintattico indipendente. Attraverso l’esame e il confronto delle caratteristiche sintattiche delle costruzioni nominalizzate indipendenti nelle due varietà si porrà l’accento sui meccanismi di sviluppo che portano il medesimo processo lungo traiettorie simili, ma non perfettamente coincidenti, cercando di carpirne un’eventuale direzionalità.

Parole chiave: nominalizzazione, insubordinazione, sudarabico moderno, Mehri, Baḥari

1. Introduzione

Il mio contributo ‘Processi di insubordinazione in sudarabico moderno’ presentato in occasione del workshop SLI ‘Tipologia e sociolinguistica: verso uno approccio integrato allo studio della variazione’ ha cercato di inquadrare in prospettiva tipologica un caso poco noto al di fuori degli studi semitici, ovvero la presenza di frasi formalmente simili a una relativa ma utilizzate come frasi indipendenti senza un antecedente espresso, altrimenti dette di frasi ‘pseudorelative’ (Pennacchietti 2007) in sudarabico moderno (d’ora in avanti SAM). Esse sono introdotte da un elemento ‘relativo’ *d* in Mehri e *l*- in Baḥari, aventi proprietà sintattiche simili ma non perfettamente coincidenti. L’intervento verteva sull’idea che questo fenomeno fosse interpretabile come un caso di insubordinazione, ovvero d’uso di una

proposizione che possiede caratteristiche formali proprie di una proposizione subordinata in contesto indipendente (Evans 2007: 367).

Si osservi l'esempio in (1):

- (1) Mehri¹
- | | | |
|-----------------------|----------------------|-------------------|
| <i>wkō</i> | <i>ḍa=ḡrib-š</i> | <i>ḍa=hōh</i> |
| come | NMLZ=sapere\PFV-2S.F | NMLZ=1S |
| <i>ḍa=a-ḡawlak</i> | <i>man</i> | <i>ḥa=ybayt=i</i> |
| NMLZ=1S-guardare\IPFV | da | DET=cammella =1S |
- ‘Come facevi a sapere che stavo cercando la mia cammella?’

Senza addentrarci nella disamina dei valori specifici delle tre occorrenze di *ḍ*- in (1), è sufficiente osservare l'apparente valore completivo della seconda occorrenza (*ḍa=hōh*) per intuire superficialmente una certa multifunzionalità che eccede il semplice uso relativo, così come le altre due occorrenze non sono istintivamente intelligibili nella loro funzione. A complicare le cose, abbiamo anche altri usi che eccedono l'uso relativo in senso stretto, come le costruzioni attributive aggettivali e genitive, due tipi di costruzioni che presentano in genere un certo grado di correlazione nelle lingue del mondo (Gil 2013):

- (2) Mehri
- | | |
|---------------------------|--------------------|
| <i>a. a=nhōr</i> | <i>ḍ=arbáy</i> |
| DET=giorno | NMLZ=quarto |
| ‘il quarto giorno’ | |
| <i>b. a=ḡáyḡ</i> | <i>ḍ=a=ḡəḡənōt</i> |
| DET=uomo | NMLZ=DET=donna |
| ‘il marito della ragazza’ | |

Avendo stabilito che fosse improprio definire *ḍ*- una semplice marca relativa e volendo evitare definizioni *ad hoc* per ogni singolo uso funzionale differente, ho continuato a lavorare sull'ipotesi di una più ampia e generica classificazione come marca associativa, cercando di integrare la letteratura esistente in merito. Nella tradizione degli studi, difatti, questa prospettiva è pressoché stata l'unica presa in considerazione, nonostante le evidenti criticità che saranno espone nel

¹ Il sistema di trascrizione impiegato è fonologico e segue le consuetudini comunemente impiegate negli studi delle lingue semitiche. Un trattino sotto il carattere indica il modo di articolazione fricativo (<*ḍ*> = [ð]), mentre il puntino indica le cosiddette ‘enfatiche’, foneticamente realizzate come eiettive o faringalizzate (<*š*> = [sʰ]~[sʕ]). Si veda Gasparini (2017) per maggiori dettagli sull'articolazione delle enfatiche.

corso di questo contributo. Un vero punto di svolta in questo apparente vicolo cieco è stato fornito durante un mio intervento all'interno del Seminario di Semitistica della Libera Università di Berlino grazie alle osservazioni di Dr. Grace J. Park, la quale mi ha suggerito di provare ad applicare il modello della nominalizzazione indipendente a questo caso di studio. Intendiamo la nominalizzazione come una 'transcategorial operation, one which derives nominals from non-nominals' (Comrie & Thompson 2007: 334; Grunow-Hårsta 2011: 216). La nominalizzazione è quindi un'etichetta inclusiva e versatile. Con *nominalizzazione indipendente* ('stand-alone nominalization') si intende la nominalizzazione di una intera frase usata sintatticamente come una principale (Yap et al. 2011a) e può essere intesa come un caso di insubordinazione in senso lato (Evans & Watanabe 2016: 1). Secondo Yap et al. (2011b: 3-8), la nominalizzazione opera su diversi livelli: argomentale vs. eventuale, lessicale vs. frasale e dipendente vs. indipendente. Accennerò ai processi di nominalizzazione all'interno del SAM in generale in quanto necessario per l'inquadramento teorico del contributo, per poi concentrarmi sulla *nominalizzazione eventuale frasale indipendente*.

La prospettiva della nominalizzazione non è stata ancora applicata al campo degli studi semitici in generale, ma è già stata proposta per lo studio di elementi morfologici di singole lingue, come nel caso di *ʔašer* e *kī* in ebraico biblico (Park 2015, 2016) nonché recentemente nel caso della nominalizzazione a livello lessicale in Soqotri, altra varietà SAM (Shibatani & Bin Makhashen 2019). Lo studio della nominalizzazione indipendente è comunque inedito nel campo degli studi sul SAM, in quanto il lavoro di ricerca su queste varietà si è incentrato principalmente su fonologia e morfologia (p. es. Lonnet & Simeone-Senelle 1997; Simeone-Senelle 2011; Dufour 2016), mentre lo studio della sintassi e dei fenomeni del discorso in generale è ancora relativamente carente (con alcune importanti eccezioni, p. es. Wagner 1953; Watson 2012) o comunque ha seguito altri approcci di indagine; alcune varietà inoltre non sono state ancora adeguatamente descritte. Solo il Mehri possiede varie descrizioni grammaticali complete (Rubin 2010, 2018; Watson 2012; Watson et al. 2020); per il Baḥari vedi Gasparini & Morris (in prep.).

Oltre al puro interesse tipologico e descrittivo, vari interrogativi muovono questa ricerca. Basandomi su due varietà SAM, ovvero

Mehri e Baṭḥari, cercherò di investigare possibili similitudini e differenze nelle strutture in oggetto tenendone in considerazione i forti legami genetici, tipologici e sociolinguistici che connettono le due varietà. Si porrà inoltre l'accento sui meccanismi di sviluppo che portano il medesimo processo lungo traiettorie simili, ma non perfettamente coincidenti.

In questo contributo considererò gli elementi *d/l-* dei *nominalizzatori*. Nel paragrafo §2 descriveremo le varietà oggetto di questo studio e i loro principali tratti tipologici. In §3 provvederemo un inquadramento tipologico della questione della nominalizzazione e delle lingue semitiche. Motiverò la mia scelta analitica riprendendo in considerazione i precedenti lavori che si sono occupati del tema (Simeone-Senelle 2003; Watson 2009; Pennacchietti 2007; Kapeliuk 2018; Rubin 2018). In §4 analizzerò nel dettaglio il fenomeno della nominalizzazione indipendente da una prospettiva sintattica e testuale. In §5, infine, proporrò alcune considerazioni conclusive sui dati qui presentati.

2. *Mehri e Baṭḥari*

Mehri e Baṭḥari sono due delle sei varietà conosciute collettivamente come lingue sudarabiche moderne, a loro volta facenti parti della famiglia semitica (afroasiatico). Il termine 'sudarabico moderno' è fuorviante in quanto queste varietà non sono in un rapporto di derivazione diretta né con le varietà arabe né con le cosiddette lingue sudarabiche antiche (o epigrafiche): parliamo in realtà di varietà dai tratti tipologici ben distinti e riconoscibili parlate da comunità per lo più semi-nomadiche nella parte meridionale della penisola arabica, all'incirca a cavallo tra la provincia yemenita orientale del Hadramawt (e l'isola di Soqatra), la provincia del Dhofar nell'Oman occidentale e le aree confinanti dell'Arabia Saudita. Queste varietà prive di riconoscimento ufficiale e a carattere principalmente o esclusivamente orale godono di differenti livelli di prestigio: per il Mehri, la varietà più diffusa, abbiamo una stima molto variabile che va dai 100.000 ai 180.000 parlanti non avendo censimenti ufficiali (Watson 2012: 1). Nonostante il contesto dominato dall'inarrestabile spinta omologante dell'arabo (dialettale e standard), il Mehri è salvaguardato dal crescente interesse delle comunità locali nel preservare le proprie identità

tradizionali; per il Baḥari, invece, abbiamo soltanto una dozzina di parlanti anziani, mentre il resto della comunità ha ormai completato il processo di assimilazione linguistica verso l'arabo e le altre varietà SAM come conseguenza dello stigma sociale che tradizionalmente affliggeva questa tribù.

Dal punto di vista sociolinguistico, il Mehri presenta variazioni sugli assi diastratico, dipendentemente dall'affiliazione tribale, e diatopico, con varietà yemenite occidentali (*Mehriyat*) e orientali (*Mahriyot*) e omanite (*Mehreyyet*). Il Baḥari viene spesso considerato una varietà tribale del Mehri – soprattutto dai parlanti di quest'ultima, a onor del vero –, nonostante fra le due lingue non vi sia completa intelligibilità. Questo giudizio è tuttavia influenzato dai delicati equilibri di potere tra le varie tribù tradizionali. Va ricordato che il Dhofar è rimasto praticamente isolato dal resto del mondo fino agli anni '70, quando è avvenuta l'unificazione del Paese; prima di allora, i suoi abitanti conducevano uno stile di vita semi-nomadico in un contesto tribale, dove lo Stato centrale era praticamente assente. Le migrazioni e le invasioni dei Janayba dal nord-est e dei Mahra provenienti dallo Yemen hanno relegato i Baḥira nella zona costiera dell'estremo est del Dhofar, vicino al confine con il governatorato di al-Wuṣṭā, dove tuttora vivono. Sconfitti dalle altre tribù, persero il controllo della loro terra e vennero assoggettati. Dopo l'unificazione dell'Oman e il prepotente ingresso della modernità, i Baḥira hanno attuato un rapido processo di deriva linguistica, cercando di eliminare lo stigma sociale cui la loro cultura tradizionale (per cui anche la lingua) erano collegati e assimilandosi ai gruppi dominanti.

Non è questa la sede per disquisire sull'eterna diatriba 'dialetto' contro 'lingua'; possiamo comunque affermare che Mehri e Baḥari mostrano forti affinità strutturali che, considerate anche le altre varietà SAM, suggeriscono l'esistenza di un continuum linguistico comprendente anche altre varietà, come Ḥarsusi e Hobyot. Il relativo isolamento sociale in cui i Baḥira hanno vissuto per un tempo imprecisato, ma comunque consistente, ha fatto sì che, con il passare del tempo, questa varietà sia venuta a sviluppare alcuni tratti divergenti rispetto al Mehri. L'esame della nominalizzazione indipendente in queste due varietà assume allora anche una rilevanza sociolinguistica: noteremo infatti aspetti di variazione relativi a un medesimo processo

operante in due varietà strettamente correlate e dalle medesime caratteristiche tipologiche.

I dati su cui è basato questo contributo provengono principalmente da Rubin (2018), Morris (in prep.) e da appunti personali tratti dal mio lavoro sul campo. Per quanto riguarda la rappresentazione della variazione dialettale del Mehri, ho esaminato principalmente dati relativi al Mehreyyet: in base al materiale disponibile non sembra esserci variazione interna al Mehri nelle proprietà della nominalizzazione.

2.1 Tratti tipologici

Il sudarabico moderno nel suo insieme possiede dei un ordine dei costituenti di tipo SVO in frasi non marcate, con un uso significativo di strutture VS(O)/VOS quando i referenti sono stati precedentemente introdotti (Watson 2012: 256-263), con testa iniziale e allineamento nominativo-accusativo. I nomi, i pronomi personali e i verbi sono marcati per genere (maschile e femminile) e numero (singolare, duale e plurale). La definitezza nominale è marcata morfologicamente da un elemento proclitico.

Come è tipico nelle lingue semitiche, il sudarabico moderno presenta morfologia non concatenativa, combinata ad affissazione flessiva e derivazionale. La marcatura TMA (Tempo, Modo e Aspetto) nella morfologia verbale presenta una distinzione fondamentale tra aspetto perfettivo (coniugazione a suffissi) e aspetto imperfettivo (coniugazione a prefissi). Per quanto riguarda la modalità, distinguiamo al solo aspetto imperfettivo i modi indicativo, congiuntivo, condizionale e imperativo, mentre il tempo è inferibile dal contesto oppure espresso tramite costruzioni analitiche. Esistono due forme verbali nominalizzate: il participio futuro (detto anche participio attivo, vedi Rubin 2007), che viene utilizzato principalmente per esprimere futuro o intenzionalità, e il sostantivo verbale.

3. *Tra relativizzazione, attribuzione e nominalizzazione in SAM*

Volendo riassumere le funzioni di *d/l-* nei termini più generali possibili, essi sono utilizzati nell'espressione dell'attribuzione in senso lato. La dipendenza di un'espressione attributiva, sia essa di natura sintagmatica nominale, preposizionale o frasale, viene marcata in SAM – così come comunemente nelle lingue semitiche – tramite due

strategie principali: o tramite semplice apposizione dell'espressione attributiva alla sua testa come nel caso di (3) e (4) o tramite l'utilizzo di una marca esplicita (Mehri *d-*, Baḥari *l-*), come in (5) e (6):

- (3) Mehri
bər ġayt
 figlio zia
 'cugino materno'
- (4) Baḥari
šar ə=rawnə ləbədāt=ha ə=yawm
 erba DET=mare seccare\PFV-3S.F=3S.M DET=sole
 'un'alga che il sole ha seccato'
- (5) Mehri
a=ğayg d=a=ğəgənōt
 DET=uomo NMLZ=DET=donna
 'il marito della ragazza'
- (6) Baḥari
yə-ğšār a=šaxəb il
 3M-diminuire\IPFV DET=dolore NMLZ
b=a=šayn=ək mən ə=dəwi
 in=DET=occhio=2S.M da DET=medicina
 'il dolore nel tuo occhio si allevia grazie alla medicina'

L'esempio di 'stato costruito' in (3) non mostra alcuna marca di dipendenza al di fuori della giustapposizione. In (4) la funzione sintattica di *šar* come complemento oggetto del verbo *ləbədāt* all'interno dell'espressione attributiva è segnalato dal pronome personale suffisso 'ritornante'. La giustapposizione è ristretta e opzionale nell'espressione di una relazione inalienabile o tra elementi non definiti²; più comuni sono invece le espressioni analitiche in (5) e (6), in cui l'attribuzione della testa definita è espressa tramite quella che per ora chiameremo 'marca associativa', *d-* per il Mehri e *l-* per il Baḥari. Queste derivano storicamente dai determinativi-relativi proto-semitici **d* e **l* le cui tracce sono attestate nella famiglia semitica in generale (Pennacchietti 2005; Huehnergard 2006; Huehnergard & Pat-El 2018); soltanto la prima è però utilizzata come fonte per la creazione della marca associativa in tutte le altre lingue sudarabiche – insieme alla variante de-fricativizzata *d-* (Simeone-Senelle 2003: 40) –, mentre la forma del

² Si veda Watson (2009) per una disamina sullo stato costruito in Mehryōt.

Baṭḥari diverge, probabilmente in seguito all'influenza areale delle varietà arabe parlate localmente o, meno realisticamente, per sviluppi diacronici indipendenti in stadi non documentati – ipotesi per cui non possiamo né potremo mai avere prove.

Al di là della divergenza nella base fonomorfológica, le molteplici funzioni di questi due elementi risultano in larga parte coincidenti, con alcune importanti differenze che saranno discusse nei paragrafi successivi e che costituiscono il punto centrale di questo contributo. Passeremo prima in rassegna gli studi in merito per poi sviluppare un modello analitico che permetta di rendere conto tanto degli usi canonici quanto degli usi insubordinati in modo coerente.

3.1 Tradizione degli studi

La bibliografia sul tema non è particolarmente abbondante. La marca associativa fino a ora è stata descritta in maniera più sistematica per il solo Mehri, mentre nel caso del Baṭḥari non esiste alcun dato pubblicato in merito, a parte qualche accenno sbrigativo in Gasparini (2018). In ogni caso tratterò i due elementi come un'unica entità nelle considerazioni che seguiranno la disamina della bibliografia, non avendo riscontrato differenze sostanziali nel loro comportamento sintattico in generale. I punti di divergenza verranno successivamente esaminati in §4.

Simeone-Senelle (2003: 239) evidenzia 'une polyfonctionnalité qui lui confère une portée syntaxique très étendue': possiamo infatti analizzarlo come 'copule, comme déictique déterminant du nom, comme relateur, connectif et relatif, joncteur de proposition complétive et comme marqueur aspecto-temporel et modal' (Simeone-Senelle 2003: 241), a seconda del contesto.

Una definizione più precisa viene data da Watson (2009: 238), la quale la descrive come 'a nominalising particle that can head an attribute or a verbal predicate'. Watson ha quindi il merito di introdurre il concetto di 'particella nominalizzatrice', ma lo applica solamente allo studio delle strutture genitivali nominali in Mahryōt. In Watson (2012: 133) divide gli usi in particella attributiva, complementatore, congiunzione in una proposizione avverbiale e modificatore verbale.

Rubin (2018: 187), sulla scia di Johnstone (1975), si orienta verso un approccio simile: '[p]erfect, imperfect, and subjunctive verbs can all be preceded by the particle *d*'. This is to be distinguished syn-

chronically from the relative pronoun *d-* and the genitive exponent *d-*, though these all derive historically from the same source'. Questa definizione desta più di una perplessità, in quanto presuppone ben tre morfemi omofoni.

L'uso insubordinato di *d/l-* viene considerato da questi autori alla stregua di un modificatore aspettuale preverbale: ciò lascia intendere indirettamente che tale marca sia l'esito di un processo di grammaticalizzazione ormai completato e che abbia quindi una portata esclusivamente morfologica, operando direttamente all'interno del sintagma verbale, piuttosto che sintattica, e quindi a livello proposizionale. Questo punto di vista è esplicito nel passaggio precedentemente riportato di Rubin, in cui distingue chiaramente un *d-* modificatore verbale aspettuale, un *d-* pronome relativo, un *d-* esponente genitivale a cui si aggiunge anche il *d-* complementatore (Rubin 2018: 378). Queste funzioni appaiono collocabili lungo un continuum sintattico opaco, che lascia ampi spazi all'ambiguità in casi come quelli presentati in (7) e (8), proposti seguendo l'interpretazione di Rubin (2018: 188). La glossa di *d-* non è specificata volutamente:

- (7) Mehri
xəṭərāt *ǧayg* *də=yə-ǧhōm*
 volta:F uomo ?=3M-andare\IPFV
bə=hōrəm
 in=DET:strada
 'una volta c'era un uomo che camminava sulla strada [*d-*
 =relativo/circostanziale]' oppure 'una volta un uomo stava
 camminando sulla strada [*d-* = marca aspettuale]'

- (8) Mehri
ə-séni=həm *də=yə-ǧtáry-əm*
 1s-vedere\IPFV=3PL.M ?=3M-parlare\IPFV-PL
 'ho visto loro che parlavano [*d-* = relativo/ circostanziale/
 complementatore]'

In entrambi i casi la selezione di una funzione unica pare forzata, in quanto si presuppone che i possibili valori funzionali di *d-* siano da considerare strutturalmente diversi.

Postulare più processi concomitanti di grammaticalizzazione risultanti in un elemento omofono, dalla distribuzione non chiara e con evidenti casi di sovrapposizione, non può essere considerato corretto.

3.2 Attribuzione nelle lingue semitiche

Per cercare una soluzione a questo problema dobbiamo invece partire dall'idea che le categorie funzionali non debbano essere codificate necessariamente in modo discreto l'una rispetto all'altra: ci servirà allora un'etichetta capace di racchiudere tutti questi possibili usi.

Per chiarirci le idee, può essere utile osservare quali siano le caratteristiche sintattiche del corrispettivo di *d/l-* in altre lingue semitiche. Tralasciando di passare in rassegna la sterminata letteratura in merito, ci limitiamo a seguire Pat-El & Treiger (2008) e Pat-El (2020) nel constatare che l'elemento morfologico che introduce e frasi relative nelle lingue semitiche presenta una sintassi differente rispetto a, per esempio, quella del relativo indoeuropeo: se in quest'ultimo troviamo comunemente il caso morfologico concordante con la funzione del pronome all'interno della proposizione relativa stessa, mentre genere e numero concordano con l'antecedente, nelle lingue semitiche la marca associativa in funzione di relativo desume tutti i tratti flessivi dalle caratteristiche del referente all'interno della proposizione matrice (Pat-El & Treiger 2008: 275).

Pennacchietti utilizza l'etichetta di 'pronomi determinativo' (Pennacchietti 1968, 1996, 2007), la quale ha il pregio di 'evitare l'equivoco della secondarietà dell'impiego genitivale rispetto a quello relativo' (Pennacchietti 1968: 63). Pennacchietti riconosce l'assenza di un valore semantico nel pronome determinativo che ne impedisce l'autonomia, vincolandolo ad una funzione di determinazione del sintagma di cui è testa sintattica, a prescindere dalla classe di parole di appartenenza del modificatore. Il medesimo autore, infine, nota la peculiare distribuzione ed estensione delle funzioni del pronome determinativo nelle lingue semitiche: esso infatti pare essere sintatticamente incompatibile con lo sviluppo dell'articolo determinativo proclitico in Nordarabico e Cananico, dove difatti ha subito una drastica riduzione degli usi o è sparito completamente, mentre le altre lingue semitiche conservano un più ampio spettro funzionale (Pennacchietti 2005).

In altre parole, l'elemento morfologico che introduce la frase attributiva/relativa nelle lingue semitiche non svolge un ruolo argomentale vero e proprio all'interno della proposizione stessa, in quanto non prende parte alla struttura sintattica dell'elemento proposizionale che introduce, fungendo solo come marca puramente associativa indicante il punto di inizio dell'espressione (Pat-El & Treiger 2008: 268, 275;

Pat-El 2020). Goldenberg (1995) e Pat-El & Treiger (2008) suggeriscono che la funzione originaria fosse allora di pronomi *sostantivante* in quanto, trovandosi l'elemento seguente giustapposto, lo governa a prescindere dal suo status sintagmatico conferendogli proprietà nominali.

3.3 Nominalizzazione in SAM

In SAM l'uso sostantivante di *d/l-* sembra avere conservato la sua potenza funzionale in molteplici ambiti sintattici: per questo credo sia necessario abbandonare la lettura tradizionale negli studi delle lingue sudarabiche di *d/l-* come relativo. Per chiarire ulteriormente questo punto prendiamo in esame un'apparente peculiarità sintattica del Mehri. Citando Rubin (2018: 81-82), quando il pronome relativo (sic) è il predicato in una proposizione non verbale con un soggetto pronominale, allora il verbo nella proposizione relativa concorda con quel soggetto pronominale, e non con il pronome relativo, come in (9b). Ciò vale anche nel caso in cui vi sia un predicativo del soggetto espresso come in (9c) oppure nel caso in cui il referente svolga la funzione di complemento oggetto nella 'relativa', come in (9d):

- (9) Mehri
- a. *də=kás-k* *t=əh*
 NMLZ=trovare\PFV-1S ACC=3S.M
 'l'ho trovato'
- b. *hōh* *də=kás-k* *t=əh*
 1S NMLZ=trovare\PFV-1S ACC=3S.M
 'sono quello che l'ha trovato'
- c. *hōh* *ǧayg* *də=xalás-k*
 1S uomo NMLZ=perdere\PFV-1S
hōrəm
 DET:strada
 'sono un uomo che ha perso la strada'
- d. *hōh* *sənnáwrət* *əd=kōn-ək* *t=i*
 1S gatto:F NMLZ=crescere\PFV-2S.M ACC=1S
 'sono il gatto che hai cresciuto (let. tu hai cresciuto me)'

Vediamo in (9a) un caso di nominalizzazione indipendente, fenomeno di cui tratteremo ampiamente in §4. Limitiamoci a dire, per ora, che (9a) costituisce un caso di nominalizzazione di un verbo dinamico nel contesto di una proposizione principale. Un raffronto tra (9a) e

gli altri esempi dovrebbe rendere palese che non ci troviamo di fronte a delle costruzioni relative, bensì a diversi casi di nominalizzazione: *dā-kásk tab* in (9b) è una proposizione nominalizzata che riveste il ruolo di predicato nominale, mentre le nominalizzazioni in (9c) e (9d) hanno entrambe un valore appositivo, reso per corrispondenza nell'uso con una relativa in traduzione italiana. In termini più banali ma forse meno oscuri, volendo trovare un corrispettivo più letterale per (9c), a scampo di equivoci, avremmo una costruzione participiale del tipo 'sono un uomo avente perso la strada'. In (9d) il ruolo di complemento oggetto del referente è richiamato dal complemento oggetto di 1s *tī*: allo stesso modo di (9c), possiamo immaginare con un po' di fantasia un (improbabile) corrispettivo italiano del tipo 'sono il gatto avente tu cresciuto me'.

Prendiamo ora in considerazione un caso normalmente descritto come frase relativa senza testa in (10):

(10) Mehri

- a. *əl b=iħəm d=āwən=in*
 NEG in=3PL.M NMLZ=aiutare\PFV.3S.M=1PL
lā
 NEG
 'nessuno tra loro ci ha aiutati (lett. 'in loro non c'è quello che ci ha aiutati')
- b. *əl b=īs mōħ lā*
 NEG in=3S.F acqua NEG
 'in lei [i.e. l'area] non c'è acqua'

Il parallelismo tra (10a) e (10b) dimostra che *d-āwən-in* in (10a) occupa la stessa posizione sintattica di un normale sostantivo in veste di soggetto della frase nominale senza dover necessariamente postulare una testa non espressa. Stessa cosa si può dire per (11), dove *əlī tēwən teh* riveste il ruolo di complemento oggetto sempre con espressione obbligatoria del ruolo sintattico del referente.

(11) Baḡhari

- fšīy-ən enħá əli tēw-ən*
 cenare\PFV-1PL 1PL NMLZ mangiare\PFV-1PL
t=eh
 ACC=3S.M
 'mangiammo per cena quello che avevamo da mangiare'

Quando la testa generica in un contesto sintattico simile è invece espressa da un pronome anaforico esplicito, questa sembra assumere un valore di struttura marcata contrastiva o restrittiva, come in (12) e (13):

- (12) Mehri
amma dēk d=āmūr
 CNTR dem.far.s.m NMLZ=dire\PFV.3S.M
'hārās-k' hārūs
 sposarsi\PFV-1S sposarsi\PFV.3S.M
 'per quanto riguarda quello che ha detto 'mi sono sposato', si è sposato'
- (13) Baḥari
wə-heb ʕar dēk li
 CONN=3S.M solo DEM.FAR.S.M NMLZ
yə-štēm=sən? yə-šyōm
 3M-comprare\IPFV=3PL.F 3M-vendere\IPFV
bə-karš
 con=tallero
 'e lui, proprio quello che li ha comprati? Li rivendeva per un tallero di Maria Teresa'

In (10a) e (11) i verbi *āwənīn* e *tēwən* sono resi argomenti (nel primo caso soggetto, nel secondo complemento oggetto) grazie al processo di nominalizzazione a opera di *d/l-*. Nelle teorizzazioni più diffuse (si veda per esempio Dryer 2007) frasi simili sarebbero descritte come frasi relative senza testa, implicando l'elisione dell'elemento testa nella struttura matrice. Sposo invece la prospettiva di Shibatani (2009) e Shibatani & Bin Makhashen (2019) per cui considerando *d/l-* nominalizzatori a tutti gli effetti, non c'è più alcuna necessità di presupporre una testa pronominale elisa. Difatti, non sembra esserci alcuna differenza strutturale tra i due casi di nominalizzazione verbale di (10a) e (11) e la seguente nominalizzazione aggettivale in Miya (ciadico, afroasiatico) in (14) (Shibatani & Bin Makhashen 2019: 10 citando Dryer 2007: 197):

- (14) Miya
 a. *má rábaz*
 REL.S.F bagnato
 'quella che è bagnata'
 b. *kàba [má rábaza]*
 veste REL.S.F bagnata
 'la veste che è bagnata'

Nominalizzazione, relativizzazione ed attribuzione appaiono essere a questo punto funzioni strettamente collegate che non poche lingue del mondo esprimono tramite la medesima costruzione. Ci troviamo di fronte a un caso tutt'altro che raro o eccezionale di utilizzo della nominalizzazione in chiave relativizzante, *contra* Comrie & Thompson (2007: 378-379). Esempi di utilizzo del medesimo morfema per la codifica sia della nominalizzazione che della relativizzazione o dell'attribuzione possono essere trovati per esempio nelle lingue sino-tibetane, in giapponese e in coreano (Yap et al. 2011b: 27): possiamo spingerci ad affermare che la costruzione nominale in certi casi sia sovrapponibile alla proposizione relativa con una testa nominale a un avanzato stadio di grammaticalizzazione, di cui la relativizzazione costituisce solo una parte delle sue funzioni (Shibatani 2009). Questo sembra essere effettivamente il caso del SAM, in cui le funzioni della marca determinativa del semitico comprendono nominalizzazione, relativizzazione e attribuzione. La categoria di proposizione relativa attributiva va considerata una delle possibili funzioni esprimibili tramite la nominalizzazione, come nell'esempio seguente:

(15) Baḥari

<i>w=alḥák</i>		<i>h=eh</i>		
CONN=aiutare\PFV.3S.M		a=3S.M		
<i>l=škkábal</i>		<i>a=riyáḥ</i>		<i>a=tawát</i>
a=volgersi\SBJV.3S.M		DET=vento		DET=direzione
<i>a=gazirat</i>	<i>il</i>	<i>b=ēs</i>	<i>garēd,</i>	<i>il</i>
DET=isola:F	NMLZ	in=3S.F	cormorano\PL	NMLZ
<i>y-ḥām</i>		<i>yā-ḡtóra</i>		<i>š=eh</i>
3S.M-volere\IPFV		3S.M-parlare\SBJV		con=3S.M

'e lo aiutò a fronteggiare il vento nella direzione dell'isola dove si trovavano i cormorani, quelli con i quali voleva parlare'

Per chiarire ulteriormente, osserviamo l'esempio (16) in cui troviamo due nominalizzatori in sequenza:

(16) Baḥari

<i>maktilt-a</i>	#	<i>šamm=uh</i>		<i>šali ber</i>
parlare:PTCP- S.M	#	antenato=3S.M		PR figlio
<i>ḥzēn. beb</i>		<i>wə=ḥa=skān=ḥa</i>		
PR	3S.M	CONN=DET=gruppo\PL=3S.M		
<i>l=il=šxölēl-uw</i>				<i>b=ilḥaglām</i>
NMLZ=NMLZ=risiedere\PFV-3PL.M		in=PR		

'parlerò di # del suo antenato šAli ber Ḥazēn. Lui e i membri della sua tribù che erano stanziati in quel tempo a il-Ḥaglām'

La doppia nominalizzazione in *l=il=šxōlēluw* dimostra bene il rapporto tra nominalizzazione, relativizzazione e attribuzione. L'elemento più prossimo al verbo segnala un caso di nominalizzazione indipendente, definendo l'informazione come contestuale e ne evidenzia lo svolgersi contemporaneamente al tempo relativo della narrazione. L'elemento più a sinistra, invece, governa l'intera predicazione verbale, relativizzandola. La funzione propriamente relativa e una più ampiamente nominalizzatrice sono evidentemente rappresentate dalla stessa marca morfologica. Ora dovrebbe essere più chiaro per quale motivo tutte le occorrenze della marca associativa negli esempi riportati in questo articolo siano state glossate come NMLZ: questa categoria inclusiva permette difatti di descrivere coerentemente tutti gli usi della marca morfologica. Non ci costringe inoltre a utilizzare glosse diverse a seconda dell'occorrenza. Contravvenendo all'associazione univoca tra morfema e glossa si produrrebbe altrimenti una contraddizione analitica.

Secondo Yap et al. (2011b: 18), nominalizzatori e dimostrativi (nella loro funzione pronominale) possiedono entrambi una funzione 'reificante', poiché creano un'entità ontologica da un evento o stato, permettendo ai parlanti di commentarlo in virtù della loro natura determinativa: per questo i dimostrativi vengono spesso grammaticalizzati in elementi discorsivi che indicano l'attitudine del parlante in relazione alla struttura narrativa e alla funzione pragmatica e testuale.

Riassumendo, abbiamo stabilito la natura intrinseca di nominalizzatore di *d/l-*, dalla quale derivano diacronicamente le altre funzioni, sia essa attributiva, relativa o dimostrativa. Questa analisi servirà a spiegare gli usi insubordinati che tratteremo nel prossimo paragrafo.

4. *La nominalizzazione indipendente*

4.1 Nominalizzazione indipendente e TMA

In questa prima parte descriverò le proprietà TMA del verbo nominalizzato indipendente, mettendo in luce gli aspetti che distinguono Mehri e Baḥari. Nella seconda, difatti, la nominalizzazione del verbo avviene necessariamente con il verbo al perfettivo, mentre in Mehri possono essere nominalizzati verbi all'aspetto perfettivo e imperfettivo e al modo congiuntivo³.

³ Ricordiamo che la distinzione perfettivo vs. imperfettivo è marcata, oltre che tramite variazione apofonica della base verbale, anche nelle desinenze della coniugazione

4.1.1 Mehri

In Mehri l'*Aktionsart* del verbo determina l'utilizzo di perfettivo o imperfettivo, a seconda del caso. Sia verbi stativi che dinamici possono essere nominalizzati alla forma del perfettivo, mentre solo i verbi dinamici lo sono all'imperfettivo. Approssimando grossolanamente, le due costruzioni esprimono rispettivamente la progressività dello stato e la concomitanza dell'evento rispetto al tempo del foreground discorsivo. La nominalizzazione del congiuntivo, che esprime un valore esclusivamente modale, è disponibile per entrambi i tipi di verbi. Rimandiamo al paragrafo §4.2 una riflessione sull'importanza della struttura informativa nella comprensione di queste costruzioni.

4.1.1.1 Perfettivo nominalizzato

In (17) e (18) i due verbi stativi nominalizzati indicano il protrarsi nel tempo della condizione predicata:

(17) Mehri

$d\partial = \acute{a}ml-\acute{a}k$ $t=i$ $l-\acute{a}k\acute{a}'$
 NMLZ=pensare\PFV-1S ACC=1S 1S-risultare\SBJV
dənyīt
 incinta
 'credo di essere incinta'

(18) Mehri

$d\partial = g\acute{a}y-\acute{a}k$
 NMLZ=avere_fame-1S
 'ho fame'

In (19) abbiamo invece un verbo dinamico al perfettivo. Questa costruzione meno frequente indica un'azione perfettiva, quindi conclusasi, i cui effetti però si ripercuotono sul presente, ovvero un valore di *anterior* (Bybee et al. 1994: 54):

(19) Mehri

$d\partial = xt\acute{a}wn$ $h\partial = b\acute{a}n=ih\acute{a}m$
 NMLZ=circoncidere\PFV.3PL.M DET=figlio\PL=3PL.M
 'hanno circonciso i loro figli'

verbale, a suffissi per il perfettivo e a prefissi o circonfissi per l'imperfettivo. Il congiuntivo invece è marcato tramite variazione apofonica della base dell'imperfettivo.

4.1.1.2 Imperfettivo nominalizzato

L'imperfettivo nominalizzato è attestato con i verbi dinamici. Viene utilizzato per indicare un evento concomitante e in corso di svolgimento rispetto a un evento o stato in foreground nella struttura informativa:

- (20) Mehri
a=nbūr=i *d=a-ṣaki=š*
 DET=giorno=1S NMLZ=1S-chiamare\IPFV=2S.F
wa=l=hama-š *t=ay* *lā*
 CONN=NEG=sentire\PFV-2S.F ACC=1S NEG
 'è tutto il giorno che ti chiamo ma non mi hai sentito' (Watson 2012: 133)

- (21) Mehri
ḥā=bū *də=γə-šḥáyk* *mān=əh*
 DET=gente NMLZ=3M-ridere\IPFV da=3S.M
 'la gente stava ridendo di lui'

4.1.1.3 Congiuntivo nominalizzato

L'utilizzo insubordinato del solo congiuntivo esprime modalità volitiva:

- (22) Mehri
abēli *γa-bōrək* *b=ūk*
 dio 3M-benedire\SBJV a=2S.M
 'che Dio ti benedica!'

Il congiuntivo nominalizzato è utilizzato per esprimere la modalità commissiva, indica cioè l'intenzionalità e l'impegno o minaccia da parte del parlante di portare a termine l'azione predicata dal verbo nominalizzato, come in (23):

- (23) Mehri
hām *əḥād* *mān=kēm* *ḵərb=áy*
 se qualcuno da=2PL.M avvicinarsi\PFV.3S.M=1S
də=lə-wbád=əh
 NMLZ=1S-sparare\SBJV=3S.M
 'se qualcuno di voi mi si avvicina giuro che gli sparero'

Rubin (2018: 193) definisce la presenza di *d-* 'idiomatic, and probably stems from the use of *d-* as a complementizer, with an implied verb of promising or swearing'. Questa affermazione è problematica: abbia-

mo solo rarissime occorrenze di simili costruzioni con la proposizione reggente espressa, per di più limitatamente a un unico verbo – *šandūr* ‘promettere’, esemplificato in (24) – mentre verbi semanticamente affini come *daxāl* ‘spergiurare’, *gəzūm* ‘promettere di astenersi da’, (*h*) *hōrəm* ‘rinunciare per rispetto a’ non sono mai seguiti da *d-* nei pochi esempi riscontrabili (Rubin 2018: 381-82):

- (24) Mehri
šandūr-k *d=əl-hāwf=ək*
 promettere\PFV-1S NMLZ=1S-pagare\SBJV=2S.M
 ‘prometto che ti pagherò’

Evans (2007: 394) evidenzia come tra le funzioni più comuni dell’insubordinazione ci sia proprio quella di indicare diversi tipi di modalità deontica, come in questo caso. Non ho ancora approfondito adeguatamente questo aspetto, quindi non ho ancora certezze in merito; inoltre, l’esiguo numero di esempi in merito rende difficile fare ulteriori elucubrazioni. Tuttavia, mi pare che la connessione tipologica tra modalità commissiva e volitiva in quanto sottocategorie della modalità deontica renda superfluo postulare una proposizione reggente sottintesa, nel caso del congiuntivo nominalizzato. È più probabile che il valore commissivo di quest’ultimo derivi concettualmente dalla ‘reificazione’ dell’enunciato volitivo del congiuntivo indipendente.

4.1.2 Baṭḥari

A differenza del Mehri, in Baṭḥari si riscontra una limitazione alle possibilità di marcatura TMA del verbo nominalizzato: a prescindere dall’*Aktionsart* del verbo troveremo invariabilmente l’uso del solo perfettivo, mentre casi di imperfettivo e congiuntivo (il quale, se usato indipendentemente, esprime comunque la modalità volitiva) non sono attestati.

La costruzione viene quindi utilizzata per esprimere senza differenziazione sia uno stato nel suo perdurare che un evento nel suo protrarsi rispetto all’evento principale. Di seguito alcuni esempi con diversi tipi di verbi: dinamico con valore perfettivo (25a), dinamico con valore abituale (25b) e stativo (25c) e (25d):

- (25) Baṭḥari
 a. el sch xaz-ōt
 NMLZ 3S.F rifiutare\PFV-3S.F

- ‘lei ha rifiutato [di sposarsi]’
- b. *a=ǧayāg* *el* *xēdam-uw*
 DET=uomo\PL NMLZ usare\PFV-3PL.M
bə=mvās
 con=rasoio
 ‘gli uomini usavano un rasoio (abituamente)’
- c. *el* *fēzʕ-ak*
 NMLZ avere_paura\PFV-1S
 ‘sono terrorizzato’
- d. *dik* *ə=maʕārikə* *dik*
 DEM.FAR.S.F DET=battaglia DEM.FAR.S.F
ber *ʕ=āho.* *kél=ko*
 STRONGPFV con=3PL.M tutto=2PL.M
el=ǧerāb-ko *t=ēs*
 NMLZ=sapere\PFV-2PL.M ACC=3S.F
 ‘quella battaglia che c’è stata tra di loro. Tutti voi la conoscete’

4.2 Nominalizzazione e periodo

In quanto proposizioni indipendenti a tutti gli effetti, le proposizioni nominalizzate del SAM possono entrare in coordinazione con un'altra frase principale o reggere una frase subordinata (Beijering et al. 2019). Esaminando l'interazione tra strutture sintattiche complesse e la semantica della combinazione di proposizioni, sul modello di Dixon & Aikhenvald (2009), notiamo ulteriori peculiarità che distinguono Mehri e Baḥari. In particolare, prenderemo in considerazione due casi ascrivibili alla macrocategoria di *addition*, ovvero *elaboration*, intesa da Dixon (2009) come un tipo di frase complessa in cui l'informazione della proposizione focale (FC='Focal Clause'), vale a dire la proposizione che fa riferimento all'attività centrale dell'enunciato, viene arricchita da una proposizione di supporto (SC='Supporting Clause') che aggiunge ulteriori dettagli al quadro informativo; e *same-event addition*, in cui le componenti della frase complessa descrivono aspetti diversi dello stesso evento, senza che un evento risulti essere chiaramente centrale. Questi rapporti semantici non sono differenziati né possiedono marche dedicate ma sono codificati in entrambe le varietà tramite il connettivo coordinante *wə* o semplice giustapposizione (Gasparini in stampa).

Da questi esempi deduciamo che in caso di *addition* il Mehri tende a marcare con la nominalizzazione tutti gli elementi verbali coinvolti.

4.2.2 Baṭḥari

Osserviamo ora invece quanto accade in Baṭḥari. Nell'esempio (28) il parlante sta iniziando un racconto. Dopo avere ricordato i nomi degli altri partecipanti all'evento nell'enunciato precedente (qui non riportato), fornisce un'informazione di background per definire il contesto in cui si verifica l'incontro con la tartaruga:

- (28) Baṭḥari
 [wə=heh bə=ḥāl héb=i el
 CONN=3S.M in=volta padre=1S NMLZ
 aǧmēd-an dar # bə=kāni
 stare_nel_pomeriggio\PFV-1PL su # in=PR
 ǧasərawwən]_{sc}. [ən-sānə ḥāmis]_{fc}
 tardo_pomeriggio 1PL-vedere\PFV tartaruga
 [bes xatəf-āt]_{sc}
 quando nuotare\PFV-3S.F
 '[e lui e il mio defunto padre, stavamo passando il pomeriggio
 su... a kāni, nel tardo pomeriggio]_{sc}. [Vedemmo una tartaruga]_{fc}
 [mentre nuotava]_{sc}'.

Nell'esempio seguente le tre proposizioni coordinate sono un altro caso di *same-event addition*, in quanto descrivono vari aspetti di un unico evento, senza una rigida struttura gerarchica nella struttura informativa. Ancora una volta soltanto la prima proposizione è nominalizzata, nonostante la ripetizione della medesima azione con identico riferimento temporale e aspettuale:

- (29) Baṭḥari
 [boh l=atxāf-k bə=tāmər],
 1S NMLZ=venire\PFV-1S con=dattero.COLL
 [atxāf-k bə=kawšərət]
 venire\PFV-1S con=quantità\DIM:F
 [w=atxāf-k bə=məšḥatāt]
 CONN=venire\PFV-1S con=animale_da_macello:F
 '[questo pomeriggio io sono venuto con dei datteri], [sono
 venuto con una modesta offerta di datteri] [e sono venuto con
 un animale da macellare]'

Un caso più complesso è presentato in (30). Per rendere più chiari i rapporti gerarchici sintattici e semantici tra le singole proposizioni introduco qui l'utilizzo di parentesi graffe che vanno a creare due blocchi di testo di supporto ('Supporting Text') che valicano i confini delle singole frasi:

- (30) Baḥari
 {[*enḥá el gīṣ-an*], [ə]=*wákt*
 IPL NMLZ avere_fame\PFV-1PL a=tempo
ḥebūr wə=baśár-ən
 DET:freddo CONN=stare_di_notte\PFV-1PL
nə-nātax]}_{ST1} {[*el ḥēbər-ən*]. [[*hes*
 IPL-pescare\IPFV NMLZ avere_freddo\PFV-1PL quando
eṭēw-ən], [*el ḥēbər-ən*],
 tornare_di_notte\PFV-1PL NMLZ avere_freddo\PFV-1PL
[nə-ntōt men ḥebūr]]}_{ST2}
 IPL-tremare\IPFV da DET:freddo
*[wə=šəwēlm-ən šeyāt]*_{FC}
 CONN=prepararsi\PFV-1PL fuoco
 '[avevamo fame], [era la stagione fredda e avevamo passato la notte a pescare con le reti]]_{ST1}. [[Avevamo freddo]. [[Tornando indietro nella notte], avevamo freddo], [tremavamo dal freddo]]_{ST2}. [Quindi ci preparammo un fuoco]_{FC}.'

In (30) l'inizio del racconto fornisce come negli altri casi una serie di informazioni che fanno da sfondo all'episodio centrale del racconto: il parlante sta narrando di una volta in cui era così stanco da essersi addormentato vicino al fuoco ed esserci caduto dentro. L'accensione del fuoco è l'evento che innesca il proseguimento degli eventi ed è l'informazione centrale attorno alla quale ruotano le informazioni circostanziali precedenti. Le informazioni di supporto sono raccolte in due nuclei distinti che ruotano intorno a due stati fisici, l'aver fame e l'aver freddo. Attorno a queste due informazioni che descrivono il contesto in cui l'evento prende luogo troviamo ulteriori elaborazioni circostanziali non marcate da nominalizzazione: soltanto le due informazioni di supporto principali vengono marcate dalla nominalizzazione, mentre gli elementi a queste coordinati no.

In (31) vediamo infine un caso di *disjunction*, in cui vengono proposte delle alternative simmetriche con uguale status. In questo caso, soltanto la prima proposizione è marcata dalla nominalizzazione

(*l-aʕāyən*), mentre le proposizioni successive non lo sono, nonostante si trovino in rapporto di coordinazione.

(31) Baḥari

[*ħad l=aʕāyən h=eh*]
 qualcuno NMLZ=fare_malocchio\PFV.3S.M a=3S.M

[*wələ ħad xadōm b=eh*]
 o qualcuno fare\PFV.3S.M in=3S.M

ʕay mən ə=ʕarr], [*ħad*]
 cosa da DET=malvagio qualcuno

aʕābb=ih] [*wələ xadōm*]
 stregare\PFV.3S.M=3S.M o fare\PFV.3S.M

b=eh səlēwə]
 a=3S.M fattura

‘[Qualcuno gli aveva messo il malocchio], [o qualcuno gli aveva fatto qualcosa di malvagio], [qualcuno lo aveva stregato] [o gli aveva fatto una fattura]’

5. Tra discorso e innovazione di sistema

5.1 Aspetto e struttura informativa

Le caratteristiche della nominalizzazione in Mehri e Baḥari sono le medesime per molti aspetti. Tuttavia, se da una parte il Mehri ha sviluppato interazioni particolari tra la nominalizzazione e la marca TMA del verbo, la nominalizzazione indipendente in Baḥari richiede necessariamente l’impiego del perfettivo.

Sembra assai probabile che l’interpretazione di marca aspettuale tradizionalmente assegnata a *d/l-* – per Simeone-Senelle (2003: 247) Mehri *d-* ‘fait partie du paradigme verbal’ – sia da mettere in relazione in primo luogo alla struttura informativa piuttosto che a un inerente valore aspettuale della costruzione stessa. Se esaminiamo i processi di organizzazione del discorso e di strutturazione dell’informazione, gli eventi codificati tramite nominalizzazione tendono a fornire informazioni di background, come del resto accade di frequente nelle lingue del mondo (Cristofaro 2016). L’impiego di costruzioni nominalizzate per l’espressione di tempo, modo e/o aspetto in ogni caso sarebbe tutt’altro che raro, come testimoniano sviluppi frequenti nelle lingue tibeto-birmane (DeLancey 2011).

Un altro punto che giudico particolarmente interessante, sia in chiave tipologica che di studio della variazione, riguarda il comportamento della nominalizzazione a livello del discorso e il suo inserirsi all'interno della struttura informativa. Anche se è necessaria una valutazione statistica significativa in tal senso, sembra di intuire che il verbo dinamico nominalizzato tenda a riflettere informazioni di background e per questo sia preferito quando, durante la narrazione, il parlante ha l'esigenza di costruire un setting per la contestualizzazione dei fatti narrati. Il verbo stativo nominalizzato invece acquisisce un valore durativo condizionato dalle dipendenze semantiche e sintattiche tra proposizioni, come è evidente negli esempi Baḥari in §4.2.2. Ne segue che questa situazione in Baḥari potrebbe riflettere un ipotetico stadio tipologico precedente rispetto al più alto livello di standardizzazione sistemica del Mehri.

5.2 L'origine della nominalizzazione

Abbiamo esaminato la questione della nominalizzazione indipendente da una prospettiva sincronica. È dunque lecito interrogarsi su quali proprietà tipologiche del SAM abbiano innescato la formazione di una struttura che non ha paralleli nelle lingue semitiche al di fuori di alcune varietà etiopiche (Kapeliuk 2018). Non possedendo alcun tipo di documentazione storica non abbiamo modo di avere prove dirette sugli eventuali sviluppi in chiave diacronica di questa struttura, ma gli studi comparativi e filologici vengono qui in nostro soccorso. Pennacchietti (2007), riprendendo Wagner (1953: 120-121), mette in relazione questa costruzione con il participio (*'relativsatz zur umschreibung des partizip'*), il quale si è completamente lessicalizzato in forme nominali in SAM. La ragione del processo andrebbe ricercata nella necessità di compensare la perdita di questa forma: la proposizione nominalizzata indipendente sopprimerrebbe quindi a un vuoto strutturale venutosi a creare nel sistema sudarabico (Kapeliuk 2018: 154) innovando l'espressione della concomitanza (vedi Cohen 1924), come peraltro già suggerito da Bittner (1913: 65-66). Saremmo quindi di fronte a un secondo ciclo di grammaticalizzazione di forme nominali, ora nominalizzate, del verbo per esprimere l'aspetto progressivo e concomitante.

5.2.1 Un indizio dalla commutazione di codice

A chiusura di questa discussione esaminiamo un caso di commutazione di codice tra Baḥari e varietà locale di arabo *janaybi*, la varietà di

maggior prestigio locale e parlata da tutti i membri della tribù compresi gli anziani.

Vediamo nell'esempio (32) che il parlante riformula e corregge la parte conclusiva dell'enunciato. Il loco coinvolto nella riformulazione è particolarmente interessante perché contiene proprio un verbo nominalizzato che in questo caso funge da modificatore nominale. Nel passaggio all'arabo il parlante riformula il verbo nominalizzato in una forma participiale attiva, caratterizzata per definizione da proprietà sia verbali (valore aspettuale progressivo) che aggettivali (funzione di attributo, accordo morfologico di tipo nominale):

- (32) Baḥari
wə=hes *nētaḥ* *warx*
 CONN=quando scendere\PFV.3S.M PR
ḡabōr *dənāwəg*
 incontrare\PFV.3S.M barca\PL
l=assēfər-nə # *[rāyḥ-āt]^{AD}* *Ḥand*
 NMLZ=viaggiare\PFV-3PL.F andare\DET-PL.F PR
 'e quando andò giù a Warx incontrò delle barche che stavano
 andando... stavano andando in India'

Credo che questo caso sia decisamente rilevante, per quanto un *hapax* all'interno del corpus di Baḥari che ho utilizzato. Il lettore potrà obiettare che una sola occorrenza non sia per nulla rappresentativa e che quindi non debba consentire generalizzazioni; tuttavia, dobbiamo considerare due variabili: la natura del corpus stesso e gli intenti dei partecipanti alla conversazione. Durante la raccolta dati la commutazione di codice è infatti stata scoraggiata sistematicamente; di conseguenza, il fenomeno si incontra soltanto accidentalmente nelle registrazioni e riguarda in genere elementi nominali.

6. Conclusioni

Con questo contributo ho cercato di mettere in discussione un punto problematico nella descrizione del sudarabico moderno tramite gli strumenti della tipologia. Da una parte ho dimostrato come all'interno di due varietà strettamente connesse un medesimo processo possa svilupparsi lungo binari per larga parte paralleli, ma con importanti elementi di discrepanza. Sicuramente le caratteristiche tipologiche del SAM hanno costituito un terreno fertile per lo sviluppo del fenomeno; rimane da

chiedersi se una delle due varietà rappresenti uno stadio più avanzato nel processo di mutamento linguistico – ovvero se uno scenario implica come passaggio l'altro – oppure se siano da considerare sviluppi paralleli. Il livello di specificità nella codifica delle informazioni TMA presente in Mehri, unito alla presenza di maggiori vincoli discorsivi nelle costruzioni complesse sembra fare propendere verso la prima soluzione: il Mehri rappresenterebbe allora uno stadio più avanzato di assegnazione di corrispondenze costruzione-funzione rispetto al Baṭḥari.

Ho inoltre cercato di fare risaltare l'importanza dell'utilizzo di modelli descrittivi adeguati da parte dei linguisti che lavorano sul campo. Il modo in cui vengono analizzati e presentati i dati ha inevitabilmente una ricaduta sul lavoro del tipologo, che di solito non può fare altro che fidarsi di quanto scritto da altri – soprattutto nel caso di lingue poco studiate e minoritarie come quelle oggetto di questo studio.

Purtroppo non avremo modo di vedere quale sarà lo sviluppo diacronico della nominalizzazione in Baṭḥari, visto il ridottissimo numero e l'età avanzata degli ultimi parlanti. Potremo però migliorare la nostra conoscenza sulle affinità e divergenze tra il Mehri e le altre varietà sudarabiche. Va ricordato che tradizionalmente le comunità locali hanno vissuto in stretto contatto tra di loro, con un alto tasso di multilinguismo garantito da frequentissime unioni matrimoniali intertribali, anche con membri delle comunità arabofone (Morris 2017: 11). Il potenziale dello studio di un terreno d'indagine peculiare come questo è sterminato e ancora da esplorare appieno.

Ringraziamenti

Ringrazio calorosamente Dr. Miranda J. Morris per avere condiviso con me i suoi dati e per il costante e tutt'altro che scontato supporto fornitomi lungo il mio percorso di ricerca. Ringrazio nuovamente Grace J. Park per i suoi illuminanti commenti durante un mio intervento all'interno del Seminario di Semitistica della Libera Università di Berlino: le sue osservazioni sui dati presentati in quella sede si sono rivelate decisive in fase di scrittura di questo contributo. Si ringraziano infine i due revisori anonimi i cui preziosi commenti hanno fornito importanti spunti di riflessione e migliorato grandemente la qualità finale di questo articolo.

Questa ricerca è parte del progetto “describing the Modern South Arabian Baṭḥari language”, ref. 40.20.0.007SL finanziato da Fritz Thyssen Stiftung.

Riferimenti bibliografici

- Beijering, Karin, Kaltenböck, Gunther & Sansiñena, María Sol (a cura di). 2019. *Insubordination: Central issues and open questions*. Berlino: De Gruyter.
- Bittner, Maximilian. 1913. *Studien zur Laut- und Formenlehre der Mehri-Sprache in Südarabien*. III. *Zum Pronomen und zum Numerales*. Vienna: Hölder.
- Bybee, Joan, Perkins, Revere & Pagliuca, William. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Cohen, Marcel. 1924. *Le système verbal sémitique et l'expression du temps*. Parigi: Imprimerie nationale.
- Comrie, Bernard & Thompson, Sandra A. 2007. Lexical Nominalization. In Shopen, Timothy (a cura di), *Language Typology and Syntactic Description* 3, 334-381. Cambridge: Cambridge University Press.
- Cristofaro, Sonia. 2016. Routes to insubordination: A cross-linguistic perspective. In Evans, Nicholas & Watanabe, Honoré (a cura di), *Insubordination*, 393-422. Amsterdam: John Benjamins.
- DeLancey, Scott. 2011. Finite structures from clausal nominalization in Tibeto-Burman. In Yap, Foong Ha, Grunow-Hårsta, Karen & Wrona, Janick (a cura di), *Nominalization in Asian languages: diachronic and typological perspectives*, 343-359. Amsterdam: John Benjamins.
- Dixon, Robert M. W. 2009. The Semantics of Clause Linking in Typological Perspective. In Dixon, Robert M. & Aikhenvald, Alexandra Y. (a cura di), *The semantics of clause linking: a cross-linguistic typology*, 1-55. Oxford: Oxford University Press.
- Dixon, Robert M. W. & Aikhenvald, Alexandra Y. (a cura di). 2009. *The semantics of clause linking: a crosslinguistic typology*. Oxford: Oxford University Press.
- Dryer, Matthew. S. 2007. Noun phrases. In Shopen, Timothy (a cura di), *Language Typology and Syntactic Description* 2, 151-205. Cambridge: Cambridge University Press.
- Dufour, Julien 2016. *Recherches sur le verbe sudarabique moderne*. Habilitation sous la direction de M. Gilles Authier: EPHE.
- Evans, Nicholas. 2007. Insubordination and its uses. In Nikolaeva, Irina (a cura di), *Finiteness*, 366-431. Oxford: Oxford University Press.

- Evans, Nicholas & Watanabe, Honoré. 2016. The dynamics of insubordination: An overview. In Evans, Nicholas & Watanabe, Honoré (a cura di), *Insubordination*, 1-38. Amsterdam: John Benjamins.
- Gasparini, Fabio. 2017. Phonetics of Emphatics in Baḥari. In Bettega, Simone & Gasparini, Fabio (a cura di), *Linguistic Studies in the Arabian Gulf* (numero speciale di QuadRi – Quaderno di Ricognizioni), 69-85. Torino: Università di Torino.
- Gasparini, Fabio. 2018. *The Baḥari language of Oman: towards a descriptive grammar*. Napoli: Università degli Studi di Napoli 'L'Orientale'. (Tesi di dottorato).
- Gasparini, Fabio. In stampa. Semantically unmarked clause linking in Mehri: the use of *wə-*. In Castagna, Giuliano & Edzard, Lutz (a cura di), *South Arabia: Old Issues, New Perspectives. Proceedings of the Workshop at the University of Erlangen-Nürnberg on December 19, 2019*. Wiesbaden: Harrassowitz.
- Gasparini, Fabio & Morris, Miranda J. In preparazione. *A descriptive grammar of Baḥari*.
- Gil, David. 2013. Genitives, Adjectives and Relative Clauses. In Dryer, Matthew S. & Haspelmath, Martin (a cura di), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Disponibile online a <http://wals.info/chapter/60>, Consultato il 2021-01-23.)
- Goldenberg, Gideon. 1995. Attribution in the Semitic Languages. In *Langues Orientales Anciennes: Philologie et Linguistique* 5(6). 1-20.
- Grunow-Härsta, Karen. 2011. Innovation in nominalization in Magar, a Tibeto-Burman language of Nepal. In Yap, Foong Ha, Grunow-Härsta, Karen & Wrona, Janick (a cura di), *Nominalization in Asian languages: diachronic and typological perspectives*, 215-254. Amsterdam: John Benjamins.
- Huehnergard, John. 2006. On the Etymology of the Hebrew Relative *šē*. In Fassberg, Steven E. & Hurvitz, Avi (a cura di), *Biblical Hebrew in Its Northwest Semitic Setting: Typological and Historical Perspectives*, 103-125. Gerusalemme: Hebrew University Magnes Press.
- Huehnergard, John, & Pat-El, Na'ama (2018). The origin of the Semitic relative marker. *Bulletin of the School of Oriental and African Studies* 81(2). 191-204.
- Johnstone, Thomas M. 1975. The Modern South Arabian languages. *Afroasiatic Linguistics* 1. 93-121.

- Kapeliuk, Olga. 2018. Insubordination in Modern South Arabian: A Common Isogloss with Ethio-Semitic? In Tosco, Mauro (a cura di), *Afroasiatic: Data and Perspectives*, 153-165. Amsterdam: John Benjamins.
- Lonnet, Antoine & Simeone-Senelle, Marie-Claude. 1997. La phonologie des langues sudarabiques modernes. In Kaye, Alan S. (a cura di), *Phonologies of Asia and Africa (Including the Caucasus)* 1. 337-372. Indiana: Winona Lake.
- Morris, Miranda J. 2017. Some thoughts on studying the endangered Modern South Arabian Languages. In Bendjaballah, Sabrina & Ségéral, Philippe (a cura di), *Journal of Afroasiatic Languages and Linguistics* 9(1). 9-32. Leiden: Brill.
- Morris, Miranda J. In preparazione. *A collection of Baḥari texts*.
- Park, Grace J. 2015. שֵׁר from light noun to nominalizer: toward a broader typology of clausal nominalization in biblical hebrew. *Hebrew Studies* 56. 23-48.
- Park, Grace J. 2016. Stand-Alone Nominalizations Formed with 'āšer and kī in Biblical Hebrew. *Journal of Semitic Studies* 61(1). 41-65.
- Pat-El, Na'ama. 2020. On the Alleged Unipartite Relatives in Semitic. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 170(2). 279-288.
- Pat-El, Na'ama & Treiger, Alexander. 2008. On adnominalization of prepositional phrases and adverbs in Semitic. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 158(2). 265-83.
- Pennacchiotti, Fabrizio. 1968. *Studi sui pronomi determinativi semitici*. Napoli: Istituto Orientale di Napoli.
- Pennacchiotti, Fabrizio. 1996. Il pronome determinativo *d-* e la complementazione dell'aggettivo nel neoaramaico di Urmia. *Israel Oriental Studies* 16. 71-84
- Pennacchiotti, Fabrizio. 2005. Ripercussioni sintattiche in conseguenza dell'introduzione dell'articolo determinativo proclitico in semitico. *Aula Orientalis. Revista de estudios del Próximo Oriente Antiguo* 23(1). 175-184.
- Pennacchiotti, Fabrizio. 2007. L'impiego di frasi pseudorelative con [come] verbi finiti. In Venier, Federica (a cura di), *Relative e pseudorelative tra grammatica e testo*, 133-148. Alessandria: Edizioni dell'Orso.
- Rubin, Aaron. 2007. The Mehri participle: Form, function, and evolution. *Journal of the Royal Asiatic Society* 3(17). 381-388.
- Rubin, Aaron. 2010. *The Mehri Language of Oman*. Leiden: Brill.

- Rubin, Aaron. 2018. *Omani Mehri: A New Grammar with Texts*. Leiden: Brill.
- Shibatani, Masayoshi. 2009. Elements of complex structures – where recursion isn't: The case of relativization. In Givón, Talmy & Shibatani, Masayoshi (a cura di), *Syntactic Complexity: Diachrony, Acquisition, Neuro-cognition, Evolution*, 163-198. Amsterdam: John Benjamins.
- Shibatani, Masayoshi & Bin Makhshen, Khaled A. 2019. Nominalization in Soqotri, a South Arabian language of Yemen. In Wetzels, Leo W. (a cura di), *Endangered languages: Contributions to Morphology and Morpho-syntax*, 9-31. Leiden: Brill.
- Simeone-Senelle, Marie-Claude. 2003. De quelques fonctions de *q-* dans les langues sudarabiques modernes. In Robert, Stéphane (a cura di), *Perspectives synchroniques sur la grammaticalisation: Polysémie, transcatégorialité et échelles syntaxiques*, 239-252. Louvain: Peeters.
- Simeone-Senelle, Marie-Claude. 2011. Modern South Arabian. In Weniger, Stefan, Khan, Geoffrey, Streck, Michael & Watson, Janet C.E. (a cura di), *The Semitic Languages: An International Handbook*, 1073-1113. Berlin: de Gruyter.
- Wagner, Ewald. 1953. *Syntax der Mehri-Sprache unter Berücksichtigung auch der anderen Neusüdarabischen Sprachen*. Berlin: Akademie-Verlag.
- Watson, Janet C.E. 2009. Annexion, attribution and genitives in Mahriyyōt. In Watson, Janet C.E. & Retsö, Jan (eds), *Relative Clauses and Genitive Constructions in Semitic*, 229-244. Oxford: Oxford University Press.
- Watson, Janet C.E. 2012. *The Structure of Mehri*. Wiesbaden: Harrassowitz.
- Watson, Janet C.E., al-Mahri, Abdullah, al-Mahri, Ali, al-Mahri, Bxayta M. K. & al-Mahri, Ahmed. 2020. *Ṭaḡhamk Afjāt: A Course in the Mehri of Dhofar*. Wiesbaden: Harrassowitz.
- Yap, Foong Ha, Grunow-Hårsta, Karen & Wrona, Janick (a cura di). 2011a. *Nominalization in Asian languages: diachronic and typological perspectives*. Amsterdam: John Benjamins.
- Yap, Foong Ha, Grunow-Hårsta, Karen & Wrona, Janick. 2011b. Introduction: Nominalization strategies in Asian languages. In Yap, Foong Ha, Grunow-Hårsta, Karen & Wrona, Janick (a cura di), *Nominalization in Asian languages: diachronic and typological perspectives*, 1-57. Amsterdam: John Benjamins.

VÍCTOR LARA BERMEJO

Evidential futures in Ibero-Romance

Il futuro morfologico può esprimere diverse letture semantiche in tutta l'area iberoromanza: tempo, modalità e inferenza o congettura. Eppure, sempre più autori segnalano che attualmente questo tempo è impiegato per l'evidenzialità, senza mantenere per le altre due letture. I ricercatori si sono focalizzati sull'introspezione ma mancano degli studi empirici quantitativi che mostrino il vero impiego di questo tempo in varietà non standard. Questo articolo ha lo scopo di descrivere l'uso attuale del futuro morfologico nel discorso dialettale e non colto, per scoprire fino a che punto è diventato una marca morfologica di evidenzialità in tutta la Penisola Iberica.

The morphological future can convey several semantic readings in the whole Ibero-Romance spectrum: time, modality and inference or conjecture. However, more and more authors point out that the current validity of this tense emerges for evidential nuances, leaving aside the other two readings. Researchers have mainly focused on their own introspection, but we lack quantitative empirical data that show the actual and non-standard usage of this tense. This paper aims to pinpoint the current employment of the morphological future in non-cultivated and dialect speech, in order to find to what extent it has become an evidentiality morphological marker all throughout the Iberian Peninsula.

Parole chiave: Futuro morfologico, evidenzialità, iberoromance, corpora dialettali

Keywords: Future tense, evidentiality, Ibero-Romance, dialect corpora

1. Introduction

The morphological future (MF) in the Ibero-Romance languages has been widely studied. Researchers have focused on the different readings it may have as well as its usage in (non-)cultivated speech. To summarise, the main works that deal with the MF in Spanish point out its scarce use to express time, both in the American and European varieties. Instead of the MF, future time is mainly conveyed by means

of the periphrasis *ir a* ('to go to') plus infinitive or the present indicative and, as a result, it emerges above all for modal readings, among which the conjectural one stands out (Real Academia Española 2009; Escandell Vidal 2014; Lara Bermejo 2016). In the case of Portuguese, Oliveira (1985) remarks the low incidence of the MF for time, as well as its scarce usage in non-cultivated registers. Moreover, Portuguese splits the MF into indicative and subjunctive, being the former useful mainly for conjectures. The latter, on the contrary, is quite employed and it is syntactically obligatory in adverbial and relative subordinates without antecedent. To replace the MF in indicative, Portuguese, as does Spanish, favours the periphrasis *ir* ('to go') plus infinitive or the present indicative. Galician also exhibits the competition between the MF and the strategies to which I have referred (Álvarez & Xove 2002; Freixeiro Mata 2006). Lastly, the literature about Catalan underlines the great validity that the MF with a temporal value has today, unlike the inexistence or scarcity of the periphrasis *anar a* ('to go to') plus infinitive or the present indicative to convey the same semantics (Wheeler et al. 1999; Pérez Saldanya 2002).

The problem with the investigations on the MF is the fact that they are based on very few data and, mostly, on examples that do not reflect orality or spontaneity. Likewise, the comparisons among dialects always refer to the distinction between the American and the European variety, without going more in depth into the different dialects that exist in each variety. Additionally, the likelihood for the MF to appear is also constrained to other nuances, such as the cultivated speech (Escandell Vidal 2019). In any case, all authors agree in determining the increase of the MF to express conjectures in all the Ibero-Romance spectrum, even in Catalan, albeit to a lesser extent (Badia i Margarit 1962).

In order to shed light and to show the real behaviour of the current MF, as well as its evolution during the last century, I have extracted data from a series of dialect corpora that cover all the Ibero-Romance languages and which I will describe below in the following section. I must underline that this paper does not focus on the strategies that appear to express a future event, but it is exclusively restricted to pinpointing the behaviour of the MF *per se* and to analysing its geolinguistic and semantic distribution in the whole Iberian Peninsula.

The importance of the corpora I have employed lies on two facts: the sociolinguistic type of the informant as well as the geographical diffusion of the MF. As will be explained below, the speakers' profile guarantees a real time change study, for it will provide the responses of NORM informants from several generations along the 20th century. Additionally, the cartographic drawing of these responses will make it possible to establish whether evidentiality has a spatial focus, a specific language has spread it towards its neighbouring varieties or it has appeared simultaneously in some or all of them.

2. *Corpus and methodology*

The sources from which I have taken out occurrences pertain to a number of dialect corpora that cover the last hundred years. On the one hand, I have analysed the data that can be found in the *Linguistic Atlas of the Iberian Peninsula* (ALPI), which was carried out in the first half of the 20th century. This is the first dialect work from the Ibero-Romance spectrum and its methodology consisted in the repetition on the part of rural and non-mobile informants, with a low educational background, and over fifty years old, of a series of pre-established words and sentences in their vernaculars. This method lacks spontaneity, but it is the first corpus we can rely on that shows dialect occurrences from people who were not conditioned by the standard. Likewise, ALPI's researchers only elicited one answer per sentence and locality. In the case of the MF, thirteen sentences were pre-established with a MF within. This means that, in a given town, the person that was surveyed could have produced the MF in every sentence out of the thirteen or in fewer. Eleven of out these thirteen sentences envisaged a future event; one denoted a modal non-conjectural reading, and another one had a conjectural value. As a result, the maps that I will show will provide the percentage of usage of the MF in each locality, taking into account that the maximum is eleven times (if the reading is temporal), or one in the case of both modal non-conjectural and conjectural.

On the other hand, I have extracted occurrences from other corpora carried out in the late 1990's and the early 21st century, whose data will be compared to those from the ALPI. The more recent corpora are based on the semi-conducted interview and are the following: *Corpus*

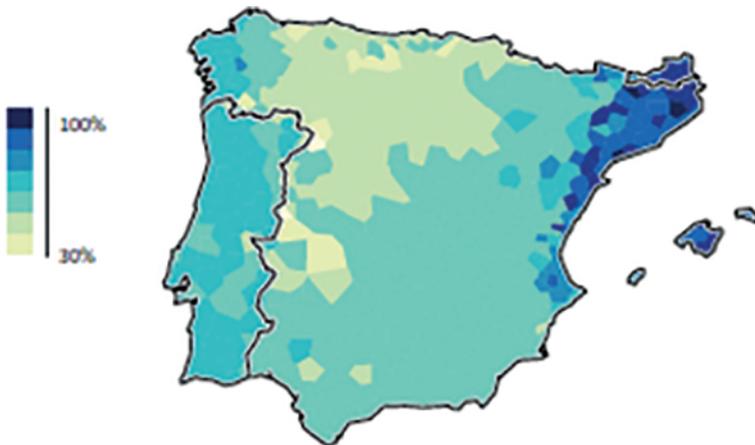
para o estudo da dialetal syntax (CORDIAL-SIN), for Portuguese; *Computerized Oral Corpus of the Galician Language* (CORILGA), for Galician; *Oral and Sound Corpus of Rural Spanish* (COSER), for Spanish; and *Corpus Oral del Català Dialectal* (COD), as well as *Corpus Dialectal del Català* (DIALCAT), for Catalan. The relevance of these corpora lies in the fact that they show oral and spontaneous speech of older informants from rural areas, without a high educational background and with little mobility; therefore, these sources guarantee dialect samples not conditioned by cultivated uses. However, not all these corpora supply the same availability of recordings. Thus, COSER provides 5,197,838 words; the CORDIAL-SIN, 600,000; the CORILGA, 29,400; and the sum of the two Catalan corpora grants 54,300 words.

Below, I will present the results on the basis of three semantic parameters, following the taxonomy that the abovementioned bibliography established for the MF: temporal MF versus non-temporal MF, and within the latter, modal non-conjectural MF versus modal conjectural MF.

3. Results

Below, I will show the results, first those belonging to the ALPI, and afterwards those from the current corpora.

Map1 - Temporal MF in the ALPI



Map 1 illustrates the frequency of the temporal MF in the ALPI. There were eleven sentences that exhibited a temporal future and, based on map 1, only Catalan in Catalonia and the Balearic Islands has systematically chosen the MF to express future time, since they provide 90%-100%.

In the rest of the Iberian Peninsula, this percentage decreases, especially in western Spanish, where 30% of the sentences were expressed through the MF, while the other 70% of informants selected another strategy, such as the present indicative or the periphrasis *ir a* plus infinitive. In any case, the incidence of the MF to convey a future time is approximately over 60% on average in all of the peninsular area, though it rises up to 80% or more in the Catalan-speaking region.

Map 2 - Modal non-conjectural MF in the ALPI



Map 2 offers the distribution of the single sentence that expressed a modal non-conjectural nuance, not associated with the conjecture. This was *The more you reap, the better*. The results show that the MF for this semantic reading was common in the Catalan from Catalonia as well as Portuguese, whereas the rest of the peninsula

does not give a single utterance of MF for this value, with a few exceptions.

Map 3 - *Modal conjectural MF in the ALPI*



Finally, Map 3 reveals the distribution of the modal conjectural MF: *He must have not arrived yet*. However, this sentence was envisaged in perfect future, but not in simple future. Furthermore, the responses from certain areas, such as Galicia or Asturias, raise doubts about how the sentence was prompted, since the informants' answers had nothing to do with a conjecture. In any case, the conjectural MF appears in all the peninsular languages, although it is hardly produced in Portuguese and, for some reason impossible to clarify, it does not appear in much of western Spanish or in Galician. It is generalised in the centre-east, as well as in Catalan, despite the fact that the literature affirmed that this possibility was scarce or non-existent in Catalan.

The data from the ALPI are scarce, but they suggest that the MF in the early 20th century emerged mostly to express time and it was infrequent to convey other readings. Below, I will present the incidence that has been found in current corpora. The results are shown in table 1.

Table 1 - *Current MF, based on the language*

	<i>Catalan</i>	<i>Spanish</i>	<i>Galician</i>	<i>Portuguese</i>
Indicative	100 (100%)	1,882 (100%)	39 (100%)	63 (14.5%)
Subjunctive	0	0	0	371 (85.5%)
TOTAL	100	1,882	39	434

As has been warned, not all corpora provide the same amount of recordings and, as a result, the divergences in number respond to this fact. Therefore, the results I will analyse have to be read taking into account this shortcoming. In any case, table 1 allows for pinpointing the current semantic validity of the MF. However, the numerical discrepancies make it necessary to count the occurrences based upon the frequency of each corpus per 10,000 words. In the case of COSER, this is 3.62073616; in CORDIAL-SIN, 7.233333333; in CORILGA, 13.265306122; finally, for Catalan, the frequency is 18,416206262.

If the results are distributed in accordance with the semantic nuances described in the corpus and methodology section, it is possible to observe that the MF depends on them, in combination with the language (table 2).

Table 2 - *Current MF, based on the semantic nuance and language*

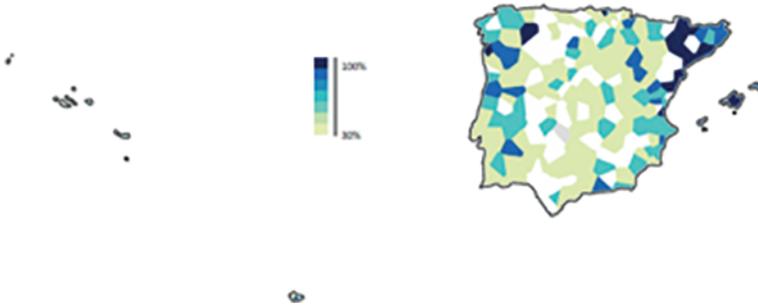
	<i>Catalan</i>	<i>Spanish</i>	<i>Galician</i>	<i>Portuguese</i>	
Temporal	84 (84%)	359 (19%)	11 (28.2%)	Ind. 7 (12.5%)	Subj. 122 (33%)
Modal conjectural	16 (16%)	1,523 (81%)	28 (71.8%)	Ind. 49 (87.5%)	Subj. 0 (0%)
Modal non-conjectural	0 (0%)	0 (0%)	0 (0%)	Ind. 0 (0%)	Subj. 247 (67%)
TOTAL	100	1,882	39	434	

Table 2 states that Catalan is the only language to exhibit the MF for expressing mainly time (84 examples with a temporal value compared to 16, that is, 84% compared to 16%). Spanish, however, offers the opposite pattern, since it presents 1,523 (81%) with a non-temporal value compared to 359 (19%) with a temporal reading. The same applies to Galician, where 11 (28.2%) occurrences with a temporal value are opposed to 28 (71.8%) without this nuance. Portuguese is the single language to possess the MF both in indicative and subjunc-

tive: the primary reading of the MF in indicative is not temporal (7 or 12.5% versus 49 or 87.5% for non-temporal nuances). Neither is it in subjunctive (247 or 67% non-temporal versus 122 or 33% temporal).

Below, I present the spatial diffusion of the MF, in line with the semantic readings foreseen in this paper: firstly, temporal MF versus non-temporal MF and, within the latter, modal non-conjectural MF versus modal conjectural MF.

Map 4 - Current temporal MF



Map 4 shows the preponderance of the current MF with a temporal value. In virtually all of Peninsular Spanish, the uses of MF with a temporal value do not go beyond 30% out of the total number of times an MF has been expressed. Moreover, in great part of the Spanish-speaking region, mainly in western Spain, no MF has been produced to convey time. Portuguese and Galician offer higher rates of temporal MF, though these are not the main values, while Catalan varies, depending on the dialect. Thus, temporal MF is hegemonic in the Balearic Islands and most of Catalonia, but the Catalan from Valencia exhibits more occurrences with a non-temporal MF than with a temporal one. In (1-6) I present some instances from all languages.

- (1) *Lo que enguany ia, segurament pues... eixe problema pues el, el tenim solucionat, segurament para este mes de octubre tindrem ia el pantano arreglat (Alicia, Valencia)*
 'It can deceive, that's the problem, but we have solved it. We are sure to have the swamp by October'
- (2) *Es nitrats és per fer créixer molt. I clar i sa gent lo que passa, clar, diven jo si sa terra... Es qui fa de pagès, i clar, no saps si s'aparsero o s'aparser, no saps si s'any venidor seguirà, entonses què fa?, que lo que va és a*

treure es profit as màxim. I i aquí ve es problema, clar (Menorca)

‘The nitrates are to grow it very much. And, well, the people say that if the soil... Who does the farmer you don’t know if he is a sharecropper and you don’t know if next year he *will still be* and then, what does he do? He wants to profit as much as he can and here is the problem’

- (3) *Y ya, cuando fuimos a llevale los regalos a los niños, nos dieron un cesto grande de caramelos y de cosas, y dice la señorita: «Se lo daremos a esta parejilla de..., que parece más formal.» (Bacares, Almería)*

‘And when we went to bring the presents to the children, we were given a basket full of sweets and the lady says: *We will give* it to this couple, which seems more formal’

- (4) *El día que vayas verás tú doña [NP] lo que te va a decí (San Francisco de Olivenza, Badajoz)*

‘The day you go, you *will see* what Mrs. [NP] is going to tell you’

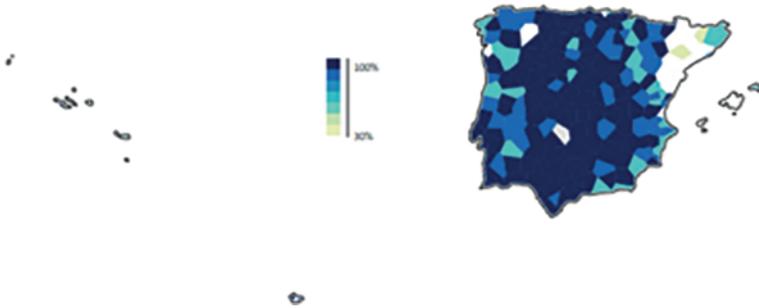
- (5) *Eu, quando eu fizer, eu vou fazer, mas vai ser com a madeira (Ponta Garça, Azores)*

‘When I *do* it, I’ll do it, but it will be with wood’

- (6) *Se vir cedo, alguma coisa há-de-se arranjar (Alvor, Faro)*

‘If it *comes* early, something must be done’

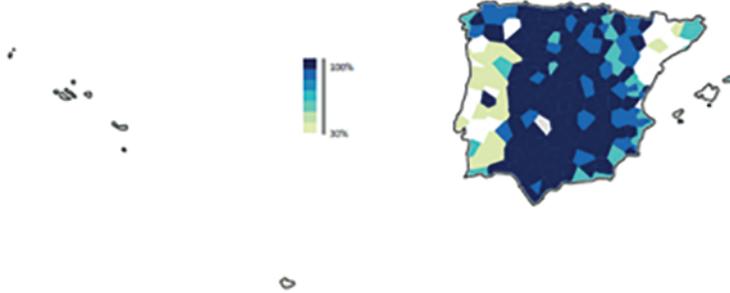
Map 5 - Current non-temporal MF



Map 5 offers the distribution of non-temporal MF. It is unanimous in most of the Spanish-speaking area, also in Portuguese and Galician and, to a lesser extent, in the Catalan from Valencia. Nevertheless, it is unusual in Catalonia and practically non-existent in the Balearic

Islands. In the following maps, I will show the incidence based upon the values conjectural and non-conjectural.

Map 6 - *Current modal conjectural MF*



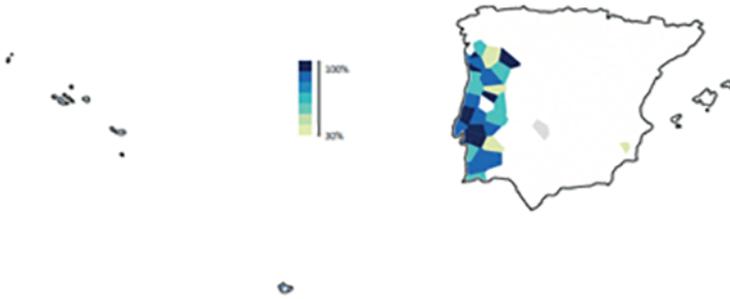
Map 6 offers the MF distribution with a modal conjectural value. The illustration is almost identical to that of map 5, but the frequency of this value decreases in Portugal. This is because the modal conjectural in Portuguese is only expressed by the indicative MF and this has slightly emerged, in contrast with the subjunctive. However, altogether, the 87.5% of indicative MF is used for conjectural reading, compared to 12.5% for a temporal one. In (7-11), I exemplify some of the occurrences of this nuance in all languages.

- (7) *Si eres municipal saberás as leyes saberás como tes que aparcar* (Monforte de Lemos, Lugo)
'If you are a policeman you *must know* the laws, you *must know* how you have to park'
- (8) *Supoño que estará no xulgado* (Matalobos, Pontevedra)
'I guess he *must be* in the court'
- (9) *Hoy tengo yo una boronia que tú no sabrás lo que es tampoco una boronia* (Jerez de la Frontera, Cádiz)
'Today I'm cooking a boronia, you *must not know* what a boronia is'
- (10) *Será verdad, claro, cuando lo dicen* (Valeria, Cuenca)
'It *must be* true, of course, if they say so'
- (11) *Lo que io sí que crec que fora de a millor de Lleida o de Girona, probablement, o no sé si la ciutat de Mallorca com ho com ho tindran, però probablement de en quant a iús de carrer i de tendes*

i de tot això pues, pues possiblement serà el que més. S'ha de dir, fora millor Girona o Lleida com a màxim, però no crec que ni Tarragona ni ni Barcelona, a lo millor (Castellón)

‘What I do believe is that outside perhaps Lleida or Girona, probably, I don’t know how it *must be* like in the city of Mallorca, but probably with regards to the use of street and shops and all that it *must be* the first one. Perhaps outside Girona or Lleida but I don’t think that in Tarragona or Barcelona’

Map 7 - Current modal non-conjectural MF



Map 7 presents the MF frequency with a modal non-conjectural value. The image depicts that this reading only occurs in Portuguese, although only if the MF inflects in subjunctive. The majority of the data in subjunctive refer to modal non-conjectural readings, but 33% of responses in this mood expresses time. In (12-13), I supply some of the occurrences in this regard.

- (12) *E depois de a pelar, fia-se também. Fia-se nas rocas também, como o linho. Só (...), não se põe (...). Não se fazem estrigas é tudo em manelos. Depois depende da lã. Se a lã for boa, a gente fia (...) melhor. Se for ruim, não se pode fiar tão boa. Mas, quando é assim para urdir, tiram-lhe mais... (Outeiro, Bragança)*

‘And after peeling it, you spin it, you spin it in the woods, as the linen. Then it depends on the wool. If the wool *is* good, the people spin it better. If it *is* bad, you cannot spin it well’

- (13) *Pronto, se for de bois é junta... Se for um boi só, é singelo. (...) E se for os dois, é uma junta (Alcochete, Setúbal)*

‘If it *is* with two oxen, it is a joint. If it *is* a single ox, it is single. And if it *is* the two of them, it is a join’

Once I have shown the data, below I will analyse them, by taking into account the diachronic evolution of the MF and the semantic evolution it has undergone during the last hundred years.

4. *Analysis*

The data described in the previous section show the evolution in the last hundred years of the usage of the MF at a dialect, oral and spontaneous level. The previous maps can be synthesised in tables 3 and 4.

Table 3 - *The MF in the early 20th century*

	<i>Portuguese</i>	<i>Galician</i>	<i>Spanish</i>	<i>Catalan</i>
Temporal	+	+	+	+
Modal (non-conjectural)	+	-	-	+
Conjectural	-	-	-	-

Table 4 - *The MF nowadays*

	<i>Portuguese</i>	<i>Galician</i>	<i>Spanish</i>	<i>Catalan</i>
Temporal	-	-	-	+
Modal (non-conjectural)	+ (subjunctive)	-	-	-
Conjectural	+ (indicative)	+	+	+/- (Valencia)

The occurrences in the ALPI indicated that the temporal value was the most widespread reading in all languages (regardless of the greater or lesser geographic incidence within each linguistic area), and it was also an established resource for the non-conjectural modality. However, the fact that the single question that was envisaged with a conjectural nuance was in perfect future as well as the fact that researchers only foresaw one sentence with such a reading, suggest that the conjectural nuance was not frequent at all one hundred years ago. In fact, the MF in perfect is cross-linguistically more inclined to convey conjectures rather than its simple version. For example, in current French, using the MF in perfect is acceptable for conjectures, but the simple future is ungrammatical for the same reading (Stage 2003). In other words: the MF in the early 20th century was above all a strategy to express time throughout the Iberian Peninsula.

The situation in table 4 is precisely the opposite to that which currently arises in the same territory. The temporal nuance is barely constructed through the MF, since it has been replaced by other strategies, such as the present indicative or the periphrasis *ir (a)* plus infinitive, with the exception of Catalan, where it continues to be hegemonic, although not as much in the Valencian area. The modal non-conjectural value is currently valid only in Portuguese, as long as it is expressed by means of the subjunctive mood. Finally, the conjectural reading has become the least marked, the most frequent in nearly all the linguistic areas, even in Catalan, whose varieties affect the likelihood for the MF to appear as conjectural. Specifically, it is the most usual reading in Valencia, but not in Catalonia and the Balearic Islands, where it is minor or non-existent.

The specialisation of the MF to mark conjecture suggests that currently this verb tense has become a morphological marker of evidentiality. According to Aikhenvald (2004), evidentiality means source of information or, rather, it deals with a grammatical category whose primary meaning is the expression of the source of information. And this is nothing more than the way in which the information has been acquired, regardless of the degree of certainty that the speaker has about the assertion or whether it is true or false.

The expression of the source of information occurs in all languages, but the development of a grammatical category for this purpose is not so frequent at all. Although a modal verb, an adverb like *apparently* or a sentence such as *I have been told* express source of information, for Aikhenvald (2004), this does not imply that a language possesses evidentiality. The author insists on the fact that evidentiality entails a grammatical category, such as morphemes or verb tenses, and that these morphemes or verb tenses arise systematically to express source of information, regardless of whether they may also be useful for other purposes. Consequently, even though the MF may denote source of information in the Romance languages, this does not mean that the MF is an evidential strategy. I am aware that other authors do not agree on this detail, but in this paper I follow Aikhenvald's theories.

I insist that the MF can only be considered as evidential, provided its primary or single meaning is source of information. Its usage as conjectures, for instance, does not turn it into an evidentiality morphological marker, since, as Aikhenvald (2004) states, a verbal tense

can also further extend a semantic reading by conveying not only time but also modality, conjecture, reportativity, etc. Nevertheless, if this semantic extension is secondary, that is, if the likelihood for the evidential nuance to appear is lower than that for time, the verbal tense cannot be labelled as evidential. In this case, the verbal tense is probably journeying towards its conversion into a morphological evidential, but it has not reached such a status.

Therefore, the MF in Catalan is not evidential, since the numbers show that the conjectural nuance is secondary compared to the temporal one (84 examples with temporal value compared to 16, that is, 84% compared to 16%). However, Spanish exhibits the opposite pattern, since it presents 1,523 (81%) with conjectural value compared to 359 (19%) with a temporal nuance. The same applies to the Galician area, since 11 (28.2%) occurrences with a temporal value are opposed to 28 (71.8%) with a conjectural nuance. Portuguese offers the identical behaviour if the MF agrees in indicative (87.5% of conjectural occurrences), but if the subjunctive is the chosen verbal mood, the dichotomy is established between temporal and modal (247 or 67% compared to 122 or 33% temporal).

Likewise, the comparison of the maps I have provided allows for determining the evolution of the MF and its conversion from tense to morphological marker of evidentiality. Whereas in Catalan, the MF is more likely to appear as temporal, the rest of linguistic areas prefer to use the MF for other values. As a matter of fact, among these three languages, the MF for time is more frequent in Galician, followed by Portuguese and Spanish in this order. Nevertheless, the expression of morphological evidentiality is usually optional or it is defective. This is the case of the Ibero-Romance, for it has created an evidential system that opposes indirect evidence that comes from inference versus the rest, which is zero marked. And it is possible to put forth that almost all of the Ibero-Romance spectrum has come up with an evidentiality system embodied in the MF because the primary meaning of this tense is to express source of information, rather than time or any other strategy. This is even the case in some of Catalan, specifically in the Catalan from Valencia, but not elsewhere throughout the Catalan-speaking region, where conjectures represent the secondary meaning of the MF.

The conversion of the MF into an evidentiality marker is not rare. According to Aikhenvald (2004), the nature of the future prompts its reanalysis as a conjecture strategy, since it conveys uncertainty. In fact, thanks to the maps shown in the previous section, it is possible to depict the different stages through which the MF has undergone from expressing time to conveying evidentiality. In the first place, the MF occurs to denote time, whereas conjectures are mainly conveyed by means of modal verbs; only every now and then, the MF can further extend another semantic nuance and also be useful for inferences: this is the case of the Catalan from the Balearic Islands.

In a second phase, the likelihood for the MF to express conjectural readings increase, but these are still minor in comparison to temporal ones; in this stage, the secondary reading is conjectural and the primary, temporal, so the conjecture continues to be a mere semantic extension of the MF, but not an evidential resource. This is the case of the Catalan from Catalonia. Subsequently, in a third stage, the conjectural occurrences outnumber the temporal ones, even though the latter are still quite frequent; in this stage, the MF is already evidential and, secondarily, it expresses time: it is the case of the Catalan from Valencia. The next phase is precisely the one witnessed in Galician, Portuguese and part of the Spanish-speaking region, where the conjectural value is nearly the unique, but tokens with a temporal value are still possible. In this phase, the MF is also evidential, but exhibits temporal behaviours as a secondary extension. Lastly, the MF disappears for temporal values and has specialised to only express source of information. This is the case of most of Spanish, mainly the west, where no example of MF to express time has been produced. In this phase, the MF as a tense has been completely replaced by other strategies, such as the present indicative or the periphrasis *ir a* plus infinitive.

5. Conclusions

The data I have analysed from different dialect corpora that comprise the last one hundred years have allowed for tracing the evolution of the MF during the last century. Based on the results, the MF was above all a resource to express future time in all linguistic areas, although it could also arise for modal readings. Currently, this situation has re-

versed and the MF is mainly employed to express source of information in Spanish, Portuguese, Galician and the Catalan from Valencia, while it is still a strategy for time in Catalonia and the Balearic Islands. This fact suggests that the MF has become a grammatical marker of evidentiality, by systematically or primarily conveying source of information (in the case of Portuguese, if it inflects in indicative).

Catalan, nonetheless, varies depending on the dialect area. Thus, Valencia is similar to the distribution of the rest of the Peninsula, since the conjectural values are more frequent than the temporal ones. This is not the case of the Catalan from Catalonia, where the conjectural MF is becoming more frequent, but it is a mere secondary extension of this tense. On the opposite extreme emerge the Balearic Islands, in which the conjectural MF is rare or non-existent, since this verb tense is used only with a temporal nuance.

The evolution of the MF follows a pattern that occurs cross-linguistically, since the nature of the future lends itself to being employed for conjectures and, therefore, to becoming an evidential strategy. Consequently, the expression of a future time in oral, dialect and spontaneous registers has been replaced by other strategies, such as the present indicative or the periphrasis *ir a* plus infinitive.

References

- Aikhenvald, Alexandra. 2004. *Evidentiality*. Oxford: Oxford University Press.
- ALPI: see García Mouton, Pilar & Fernández-Ordóñez, Inés & Heap, David & Perea, María Pilar & Saramago, João & Sousa, Xulio. 2016.
- Álvarez, Rosario & Xove, Xosé. 2002. *Gramática da lingua galega*. Vigo: Galaxia.
- Badia i Margarit, Antoni. 1962. *Gramàtica catalana*. Barcelona: Enciclopedia Catalana.
- COD: see Perea, María Pilar & Viaplana, Joan. 2008-.
- CORDIAL-SIN: see Martins, Ana Maria. 2000-.
- CORILGA: see Regueira Fernández, Xosé Luís. 2012-.
- COSER: see Fernández – Ordóñez, Inés. 2005-.
- DIALCAT: see Martí Antonin, Maria Antonia. 2006.

- Escandell Vidal, Victoria. 2014. Evidential futures: the case of Spanish. In De Brabanter, Phillippe & Kissine, Mikhail & Sharifzadeh, Saghie (eds.), *Future times, future tenses*, 221-246. Oxford: Oxford University Press.
- Escandell Vidal, Victoria. 2019. El futuro simple del español. Sistema natural frente a usos cultivados. *Verba Hispanica* 26. 17-35.
- Fernández – Ordóñez, Inés. 2005-. *Corpus Oral y Sonoro del Español Rural*. Madrid: Universidad Autónoma de Madrid. <<http://www.corpusrural.es>> [consulted: May 2019].
- Freixeiro Mata, Xosé Ramón. 2006. *Gramática da lingua galega*. Vigo: A Nosa Terra.
- García Mouton, Pilar & Fernández-Ordóñez, Inés & Heap, David & Perea, María Pilar & Saramago, João & Sousa, Xulio. 2016. *Atlas Lingüístico de la Península Ibérica (ALPI-CSIC)*. Madrid: CSIC. <<http://www.alpi.csic.es/>> [consulted: January 2019].
- Lara Bermejo, Víctor. 2016. La expresión del futuro en las lenguas romances de la Península Ibérica. *Boletín de la Real Academia Española* 96(314). 529-558.
- Martins, Ana Maria. 2000-. *Corpus Dialectal para o Estudo da Sintaxe*. Lisbon: Centro de Linguística. <<https://clul.ulisboa.pt/en/recurso/cordial-sin-syntax-oriented-corpus-portuguese-dialects>> [consulted: June 2019].
- Martí Antonin, Maria Antonia. 2006. *Corpus Dialectal del Català (DIALCAT)*. Barcelona: Universitat de Barcelona. <<http://clic.ub.edu/corpus/es/node/127>> [consulted: April 2019].
- Oliveira, Fátima. 1985. O futuro em português: alguns aspectos temporais e/ou modais. In *Actas do 1º Encontro da Associação Portuguesa de Linguística*, 353-373. Lisbon: Associação Portuguesa de Linguística.
- Perea, María Pilar & Viaplana, Joan. 2008-. *Corpus Oral del Català Dialectal*. Barcelona: Universitat de Barcelona. <<http://www.ub.edu/ccub/corpusoraldialectal-cod.html>> [consulted: April 2019].
- Pérez Saldanya, Manuel. 2002. Les relacions temporals i aspectuals. In Solà, Joan & Rigau, Gemma (eds.), *Gramàtica del català contemporani*, 2567-2662. Barcelona: Empúries.
- Real Academia Española. 2009. *Nueva gramática de la lengua española*. Madrid: Espasa.
- Regueira Fernández, Xosé Luís. 2012-. *Corpus Oral Informatizado da Lingua Galega*. Santiago de Compostela: Universidade de Santiago. <<https://ilg.usc.es/gl/node/1016>> [consulted: May 2019].

- Stage, Lilian. 2003. Les valeurs modales du futur et du présent. In Birkelund, Merete & Boysen, Gerhard & Kjaersgaard, Paul Soren (eds.), *Aspects de la modalité*, 203-216. Tübingen: Max Niemeyer Verlag.
- Wheeler, Max & Yates, Alan & Dols, Nicalau. 1999. *Catalan: a comprehensive grammar*. London: Routledge.

ANTONIETTA MARRA

Forme di futuro nello Slavo del Molise tra sociolinguistica, tipologia e linguistica del contatto

In molte lingue del mondo le forme che designano temporalità futura hanno la loro origine in strutture che esprimono un valore modale veicolato dagli elementi che, nel processo di grammaticalizzazione, assumeranno funzione ausiliaria nelle perifrasi verbali o si trasformeranno in clitici o affissi verbali. In questo contributo sono descritte alcune forme di futuro nello Slavo del Molise (SLM), una varietà croata in prolungato contatto con l'ambiente romanzo circostante. Si evidenzierà quali di queste forme sono originarie dell'area balcanica e sono comuni alle altre varietà croate e quali, invece, trovano nelle varietà romanze dell'Italia Meridionale possibili corrispondenze. Se ne esamineranno, inoltre, valori e funzioni anche in correlazione alla diffusione d'uso dello SLM nelle comunità minoritarie esaminate e nelle tipologie di parlanti intervistati.

Parole chiave: slavo molisano, forme di futuro, grammaticalizzazione, modalità, contatto slavo-romanzo.

1. *Introduzione*

Le forme di futuro appaiono particolarmente mutevoli in molte lingue del mondo. Per le lingue indeuropee si ipotizza che questa tendenza all'instabilità formale sia dovuta al processo di trasformazione che ha portato alla espressione di valori temporali subito da quelle strutture che inizialmente hanno avuto lo scopo di sostituire una supposta originaria forma sintetica indeuropea. Poco importa, tuttavia, se tutto sia conseguenza o meno della perdita di questa forma originaria, visto che il processo di trasformazione delle forme di futuro è ampiamente documentato in diverse lingue storiche che mostrano una certa ciclicità anche nella natura sintetica o analitica delle forme di espres-

sione di questa dimensione temporale (si pensi ad esempio a quanto accaduto nella storia del latino; v. oltre, par. 2.1).

Interessante appare, ai fini dell'analisi dei dati qui proposti, la dinamicità che queste forme hanno manifestato nella storia delle lingue, determinata da processi di continua e ciclica grammaticalizzazione in termini temporali di forme che originariamente veicolavano significati per lo più di tipo modale o aspettuale. Nelle definizioni delle forme di futuro, infatti, oltre al riferimento ad eventi che avranno temporalmente luogo dopo il momento dell'enunciazione, vengono richiamati anche elementi che hanno un nesso con le intenzioni e le predizioni di chi si esprime, cioè con le azioni e gli stati di cose non ancora realizzati e che sono il prodotto di una volontà (intenzioni) o che riguardano avvenimenti che sfuggono al controllo dell'agente (predizioni; cfr. Bybee *et al.* 1994: 244; Dahl 2000: 309-313). Non meraviglia, dunque, che i significati espressi dalle forme di futuro possano rientrare sia nella dimensione del *realis* che dell'*irrealis* (cfr. Bybee *et al.* 1994: 237; v. però anche Mauri & Sansò 2012 e Sansò 2020) e possano essere il frutto di mutamenti e processi di grammaticalizzazione che chiamano in causa dimensioni della modalità e dell'aspettualità (vd. tra gli altri Schwegler 1990, Bybee *et al.* 1994, Dahl 2000; Sansò 2020).

Nelle pagine dedicate al futuro, Comrie (1985: 44), a questo proposito, segnala la difficoltà di individuare nell'analisi crosslinguistica un'unica procedura di categorizzazione delle forme di futuro e conclude che solo l'analisi delle singole forme nelle diverse specifiche lingue può fornire la giusta indicazione di collocazione di queste forme nell'area della temporalità o della modalità. La difficoltà che emerge nella riflessione di Comrie (1985) è data dal fatto che sebbene, in una rappresentazione lineare del tempo, il futuro sia posto rispetto al presente allo stesso modo in cui si colloca il passato (anche se nella direzione temporale opposta), in realtà il futuro non è semplicemente ciò che si realizza oltre il momento del presente, ma accoglie in sé necessariamente un valore di minore definitezza e minore fattualità (Chung e Timberlake 1985: 206) rispetto alle forme di passato:

The future [...] is necessarily more speculative, in that any prediction we make about the future might be changed by intervening events including our own conscious intervention. Thus, in a very real sense the past is more definite than the future. Following on from this, one might argue that while the difference between past and present

is indeed one of tense, that between future on the one hand and the past and present on the other should be treated as a difference of mood rather than one of tense. (Comrie 1985: 43-44).

Questa riflessione non esclude tuttavia la possibilità che alcune forme di futuro non presentino alcuna espressione di modo e che, dunque, il futuro possa essere considerato come categoria descrittiva necessaria della temporalità per quelle lingue che, in forma originaria o per effetto di un avanzato processo di grammaticalizzazione, possiedono espressioni di futuro con valore eminentemente temporale.

Questo mio contributo ha lo scopo di descrivere alcune forme di futuro nello Slavo del Molise (SLM), lingua minoritaria croata in prolungato contatto con varietà romanze dell'Italia meridionale. La descrizione di queste forme è anticipata da una breve disamina delle strutture più diffuse per l'espressione del futuro in alcune lingue del mondo e in particolare in alcune varietà che sono vicine allo SLM genealogicamente o per condivisione di spazi geografici (nei luoghi di origine dello SLM, i Balcani, o in quelli di arrivo dopo la sua migrazione nell'Italia meridionale).

Sulla base dell'individuazione delle diverse tipologie di futuro riscontrate, sono descritte le forme di futuro nello SLM. Di queste sono evidenziate le distribuzioni in due comunità SLM, cercando di individuare eventuali correlazioni da un lato con il ruolo delle lingue romanze nelle due diverse comunità linguistiche e con quella che sinteticamente possiamo definire la vitalità linguistica (Unesco 2003) dello SLM nelle due comunità, dall'altro con alcune variabili sociodemografiche e sociolinguistiche degli informanti coinvolti.

2. Forme di futuro

Vista la difficoltà di riconoscibilità di quelle che possano essere considerate forme di futuro, è in primo luogo necessario perimetrare l'oggetto di indagine di questo lavoro.

Sono qui prese in esame quelle forme verbali che vengono considerate espressioni di una referenza temporale di futuro (d'ora in poi RTF) che possano o meno essere accompagnate da un eventuale valore modale. In particolare, l'analisi si limiterà a quelle forme che esprimono un significato di futuro rispetto al momento dell'enunciazione e che hanno in quest'ultimo l'unico momento di riferimento

(Bertinetto 1991: 18-19). Queste forme esprimono molto spesso anche ulteriori valori implicati dallo stesso valore temporale, ovvero intenzione, previsione o programmazione (Dahl 2000). Sono, questi, come si è detto, valori aggiuntivi necessariamente presenti in una forma verbale di questo tipo (Comrie 1985: 43). Le forme di futuro qui considerate rientrano nella categoria del cosiddetto *future I* (v. tra gli altri, Migdalski 2006); sono dunque escluse da questa analisi sia le forme che indicano un evento futuro precedente ad un ulteriore momento di riferimento situato nel futuro (come ad esempio il cosiddetto *futuro anteriore*, definito più coerentemente da Bertinetto (1991: 15) *futuro composto*) sia quelle di futuro nel passato.

Pur restringendo il campo delle forme di futuro da prendere in considerazione, è utile ricordare che nelle lingue del mondo è possibile individuare anche all'interno della tipologia di *future I* ulteriori distinzioni. Bybee *et al.* (1994: 244-251) evidenziano, con una ricca serie di esemplificazioni dalle lingue del loro campione, l'esistenza di forme verbali che permettono di distinguere tra futuro semplice e futuro imminente, futuro definito e indefinito, futuro atteso, previsto e pianificato. Non tutte le lingue del mondo dispongono di costruzioni altamente grammaticalizzate per esprimere queste opposizioni. In francese ad esempio è grammaticalizzata la perifrasi imminente con il verbo *aller* + infinito che trova corrispondenza con la perifrasi italiana introdotta da *stare per* + infinito (1); in inglese la forma indicativa del presente è utilizzata per esprimere eventi futuri pianificati (2), come accade anche in italiano e in alcune altre lingue (v. oltre) mentre l'uso della perifrasi *will* + infinito è utilizzata per esprimere intenzioni o predizioni.

- (1) *je vais partir*
 io vado-PRS.1SG partire-INF¹
 'sto per andare via'

- (2) *the bus leaves at noon*
 il autobus parte-PRS.3SG a mezzogiorno
 'l'autobus parte a mezzogiorno' = 'è programmato che
 l'autobus parta a mezzogiorno / l'autobus con certezza partirà
 a mezzogiorno'

¹ Sono glossate morfologicamente solo le forme che appaiono significative ai fini del nostro discorso, quindi quelle verbali.

È però difficile che la complessa articolazione delle sfumature cognitive che possono riguardare l'espressione del valore di futuro trovino tutte formalmente spazio in una stessa lingua (v. Bybee *et al.* 1994).

Dahl (2000: 310) propone di considerare come fondamentale una opposizione tra RTF basata sull'intenzionalità (*intention-based future time reference*) e RTF basata sulla predittibilità (*prediction-based future time reference*). Tuttavia, anche questa opposizione cognitivamente così saliente è di fatto poco evidente nella opposizione delle forme di RTF in quanto, afferma Dahl (2000):

[...] markers that are originally restricted to intention-based FTR [= Future Temporal Reference] tend to develop into general future markers, which include prediction-based FTR as central cases but can in the normal case still be used for intention-based FTR. In fact, whether FTR is overtly and obligatorily marked in prediction-based sentences can be used as one of the major criteria for whether it is grammaticalized in a language or not. (Dahl 2000: 310)

Processi di grammaticalizzazione di RTF di questo tipo sono molto diffusi e, come vedremo dai dati, ci guidano ad affermare che lì dove sono presenti perifrasi con ausiliari che esprimono volontà, queste rappresentano la forma basica di espressione della RTF.

Alla opposizione intenzionalità / *V*3. predittibilità, Dahl (2000: 311) aggiunge un'ulteriore categoria, quella della pianificazione (*scheduling*). Troviamo l'espressione di questa dimensione nell'esempio inglese in (2) e in molte altre lingue per frasi dello stesso tipo, nelle quali sembra che il valore temporale della forma grammaticale di presente venga neutralizzato a favore di una referenza temporale altra. In effetti, forme di presente con valore temporale altro (e più precisamente di futuro) sono utilizzate anche in frasi che indicano pianificazione immediata da parte del parlante in inglese, russo, francese, serbo, croato.

2.1 Le diverse forme grammaticali con RTF in alcune varietà dei Balcani e dell'Italia Meridionale

Seguendo l'individuazione delle forme grammaticali con RTF di Bybee *et al.* (1994) e di Dahl (2000), propongo qui di seguito una breve descrizione di alcune *famiglie di forme* (*gram families*) presenti in lingue europee dei Balcani, luogo di origine dello SLM, e della penisola italiana, territorio di arrivo della migrazione slavo-molisana.

Nell'area balcanica non esiste un futuro sintetico e molte lingue hanno costruzioni con RTF di tipo devolitivo, cioè nate a partire da un verbo con il significato di 'volere' (v. Dahl 2000: 323). Le perifrasi sono costruite in diversi modi. Per quanto riguarda le varietà più vicine allo SLM (croato e dialetti serbo-croati), il verbo *volere* ha una forma ridotta e flessa per numero e persona. Questo clitico può essere seguito da un infinito (imperfettivo o perfettivo) del verbo lessicale o da una costruzione congiuntiva con forme finite del verbo lessicale. La forma con la costruzione congiuntiva non risponde alle norme del croato standard ma di fatto è usato dai croatofoni che vivono nelle aree serbo-croate sudorientali (Tomić 2004: 519). In questo territorio sono presenti i dialetti serbi sudorientali nei quali la costruzione delle forme di futuro prevede una particella modale che ha origine dalla terza persona singolare del verbo *volere* seguita dalla struttura congiuntiva: *ja će da pišem*, lett. 'io vuole che scrivo' (Browne 1993: 330; Vaillant 1966: 109; Tomić 2004)². La particella modale *će* viene sostituita dalla forma *ću* che presenta le marche di flessione della prima persona singolare in quei contesti in cui è manifestata la volontà del parlante (Tomić 2004: 521).

Lì dove è utilizzata la forma flessa del clitico, come in croato standard, per la legge di Wackernagel secondo la quale il clitico deve occupare la seconda posizione della frase, la forma breve dell'ausiliare non deve necessariamente precedere l'infinito del verbo lessicale ma può seguirlo. L'ordine *infinito+clitico dell'ausiliare* produce una caduta della marca di infinito e una conseguente fusione dell'ausiliare con il verbo principale (3):

- (3) Croato standard
ćini-ć-u (*< ċiniti ću < ċiniti ću*)
 fare.IPFV- volere.AUX-PRS.1SG
 'farò'

Nei Balcani è presente anche una struttura con RTF costruita con l'ausiliare *avere* che originariamente aveva molto probabilmente un

² Anche in bulgaro e macedone l'espressione del futuro presenta forme perifrastiche con il verbo *volere*, sebbene in forma fissa non flessa (Vaillant 1966: 109-110; Comrie e Corbett 1993; Dahl 2000), come nei dialetti serbi sudorientali. Nelle varietà serbo-croate kajkave, invece, il futuro perifrastico è quello tipico anche dello sloveno, con *bum* (*< budem*): *imel bum* 'avrò' (lett. 'avuto sarò').

valore debitivo³. Tra le lingue balcaniche che presentano questa forma ci sono il bulgaro, la varietà ghega albanese e il rumeno dei Balcani. Mitkovska e Bužarovska (2014), inoltre, evidenziano che in macedone le strutture con la forma non flessa di *avere* esprimono un valore temporale quando c'è certezza dell'evento futuro.

Per quanto riguarda le varietà romanze, tra le lingue d'Europa sono le uniche ad avere una forma sintetica dedicata per la RTF (insieme a varie forme perifrastiche, come si è già visto nel par. 2). Le loro attuali forme di futuro sintetico sono esito di grammaticalizzazioni a partire da costruzioni costituite da *infinito del verbo lessicale + habeo* (4a).

- (4) a. *cantare habeo*
 b. *habeo cantare*
 c. *habeo ad cantare*
 d. *habeo de cantare*
 e. *venio ad cantare*
 f. *volo cantare*
 g. *debeo cantare* (cfr. Schwegler 1990: 124)

Questa sequenza di *infinito + habeo* era una delle varie forme presenti sul territorio romanzo. Come si può osservare negli esempi al punto (4), le costruzioni di partenza presentavano diversi ausiliari e differenti ordini dei costituenti. Tutte le sequenze, inoltre, esprimevano originariamente un valore modale di obbligo o necessità⁴. Tali costruzioni

³ Ho scelto di utilizzare il termine *debitivo* anziché *deontico*, certamente più noto e diffuso per descrivere questo tipo di strutture, per varie ragioni. Per la modalità deontica, indicata in letteratura come una delle modalità di base che si manifesta in opposizione a quella dinamica e epistemica (v. tra gli altri Nuyts 2016), non sempre si evidenzia un valore di necessità (oltre che di obbligo) che è invece richiamato dalla modalità dinamica (sebbene in particolari condizioni). Inoltre alla modalità deontica si fa riferimento per un valore di *permissività* (vd. sempre Nuyts, ma anche Palmer 2001) che non mi sembra sia presente in queste forme. L'uso del termine *debitivo*, utilizzato qui per indicare valori modali ma non di modo (come è invece utilizzato nella descrizione del sistema verbale del lettone; v. ad esempio Lokmane & Kalnača 2014), mi permette di includere i significati di obbligo o di necessità (entrambi presenti nelle forme di debitivo ma non in quelle deontiche) utili nell'economia del mio discorso sulle forme di futuro in SLM, senza dover affrontare in questa sede una discussione sulla complessa e non del tutto condivisa ripartizione dei valori modali (si confronti ad esempio la triplice ripartizione di Nuyts con quella di Palmer 2001) e dei parametri utilizzati per la loro definizione.

⁴ Alcuni studiosi, però, come Bybee *et al.* (1994: 261-262) parlano di *predestination*, cioè di valore modale prospettivo correlato alla espressione della volontà del soggetto e non ad una modalità debitiva o deontica.

si sono affiancate alle precedenti forme sintetiche del latino classico (*cantabo* ‘canterò’; *cantabitur* ‘sarò cantato’; *cantavero* ‘avrò cantato’) e le hanno progressivamente sostituite.

Tra le forme in (4) solo la sequenza *infinito* + *habeo*, con questo ordine degli elementi, si è gradualmente stabilizzata e la successiva erosione e grammaticalizzazione degli elementi costitutivi hanno permesso lo sviluppo di un significato temporale e la realizzazione della nuova forma sintetica di futuro che troviamo anche in italiano, oltre che in francese, spagnolo, portoghese, occitano, catalano e romancio, nelle sue diverse forme flesse (cfr. Rohlfs 1968: 331-338; Schwegler 1990: 124; Kaye 2009).

Nelle varietà romanze dell’Italia Meridionale, però, al sud della linea Viterbo-Pescara-Ancona (Rohlfs 1968: 333) il futuro non è espresso dall’esito della sequenza *infinito* + *habeo* (4a), ma dalle forme di presente indicativo. Forme di presente con funzione referenziale di futuro sono infatti molto diffuse sia nelle varietà dialettali italo-romanze, sia nell’italiano parlato (Bertinetto 1991: 69-71).

A questa forma di presente con RTF si aggiungono anche gli esiti della sequenza (4c) *habeo ad cantare*. Rohlfs (1968: 335) annota quest’ultima forma descrivendola come “una sorta di futuro in cui ancora si sottintende un poco l’idea di necessità”. La ritroviamo in Sicilia, Puglia, Lucania, Abruzzo e anche nel Molise (5)⁵.

(5)	Campano (Benevento)	<i>eggìa a purtà</i>	‘devo portare’/‘porterò’
	Molisano (Campobasso)	<i>av’ a purtà</i>	‘devi/e portare’/‘porterai/porterà’
	Abruzzese	<i>hi/a da purtà</i>	‘devi/e portare’/‘porterai/porterà’
	Pugliese (Bari)	<i>agghia a candà</i>	‘devo cantare’/‘canterò’
	Pugliese (Taranto)	<i>av’ a voni</i>	‘deve venire’/‘verrà’
	Lucano	<i>l’aggi’ a manna</i>	‘devo inviarlo’/‘lo invierò’
	Siciliano	<i>l’ai a mannari</i>	‘devo inviarlo’/‘lo invierò’

Si vedano a questo proposito anche altre forme del dialetto molisano in cui *habeo* è sostituito da *teneo*, che spesso, nelle varietà meridionali, veicola il significato di ‘avere’ (6). Anche con il verbo *teneo* la modalità debitoria è ancora fortemente presente e sono chiari i valori di obbligo e necessità con esso espressi.

⁵ Tutti gli esempi in (5) sono tratti da Rohlfs (1968: 335-336), tranne quello in campobassano, raccolto da fonte orale. Questa struttura con RTF si trova anche, al di fuori della parte meridionale della penisola, in Toscana (fiorentino popolare), in Corsica e in Sardegna.

- (6) Molisano (Jelsi - Cb)
ten-ge a ji a Cembuasce dumenaca
 tenere-PRS.1SG a andare-INF a Campobasso domenica
 ‘domenica devo andare a Campobasso’/‘domenica andrò a
 Campobasso’

3. Le forme di futuro nello SLM

Nello SLM le RTF semplice possono essere espresse con diverse forme (7):

- (7) a. presente indicativo della forma breve del verbo *tit* ‘volere’
 (< *hotit*) + infinito;
 b. presente indicativo della forma breve del verbo *imat*
 ‘avere’/ ‘dovere’ + infinito;
 c. presente indicativo;
 d. presente indicativo del verbo *stat* ‘stare’ + *za* + infinito.

Tre delle quattro forme sono costruzioni perifrastiche. Le prime tre RTF sono attestate anche in Rešetar (1997: [1911] 140-143); la quarta, riscontrata in poche occorrenze nel mio corpus (v. par.4), è segnalata anche in Breu e Piccoli (2000: 411). Sia per il numero di occorrenze ridotte, sia per lo specifico valore imminenziale che la differenzia dalle altre forme, la struttura (7d) non sarà discussa in questa sede.

Non meraviglia la compresenza di diverse forme con RTF, come sottolineano anche Bybee *et al.* (1994):

As the numbers suggest, it is not uncommon for a language to have more than one gram which has future as a use. [...] As we have argued here and elsewhere (Bybee and Pagliuca 1985, 1987; Bybee, Pagliuca, and Perkins 1991), such apparent duplication may be viewed as a consequence of the independent development of grams from distinct sources and from similar sources at different periods, which produces layers of relatively old markers underlying layers of more recently evolved ones. As we shall see, the continued viability of multiple forms in a given language is insured by differences in the range of uses to which each may be put. (Bybee *et al.* 1994: 243)

Le osservazioni di Bybee *et al.* (1994) appaiono convincenti in relazione alle forme con RTF dello SLM.

La prima forma perifrastica (7a) è una struttura riconoscibile come devolutiva (Breu e Piccoli 2000: 413; Marra 2005, 2019), data

la presenza dell'ausiliare *volere* utilizzato per costruirla. Il grado di grammaticalizzazione raggiunto da questa struttura appare tale da non esprimere necessariamente *intenzione* da parte del parlante (cfr. De Roberto 2010); appare cioè uno di quei casi in cui il processo di grammaticalizzazione porta ad una coincidenza delle forme di RTF intenzionali con quelle di "futuro generale" (cfr. par. 2, cit. di Dahl 2000). Pur tenendo conto di questo, considerando che in alcuni casi la struttura (7a) conserva il suo originario valore modale in opposizione alla forma (7b), data la presenza dell'ausiliare *tit* 'volere', si farà riferimento a questa forma con la denominazione di *futuro devolutivo* o semplicemente *futuro con tit*.

Attualmente, di norma, la perifrasi di futuro con *tit* presenta le forme brevi dell'ausiliare⁶. Solo raramente alcuni anziani utilizzano la forma piena dell'ausiliare e solo per le persone plurali, a differenza di quanto riscontrato da Rešetar (1997 [1911]: 141) a inizio secolo scorso, quando si notava una diffusa presenza delle forme piene dell'ausiliare nelle forme perifrastiche con persona plurale: in questi cento anni, dunque, si è senz'altro realizzato un avanzamento nel percorso di grammaticalizzazione della perifrasi con una progressiva cliticizzazione dell'ausiliare. Anche in SLM il verbo lessicale che segue l'ausiliare può essere sia imperfettivo sia perfettivo.

È evidente la rispondenza di questa forma di futuro con quella presente nel croato standard e in altre varietà dialettali croate (v. par. 2.1). Ricordiamo a questo proposito che lo SLM si colloca all'interno dei dialetti croati e le ipotesi sul suo arrivo sulla penisola italiana indicano un periodo tra XV e XVI secolo, periodo nel quale questa forma di futuro era già presente nell'area balcanica d'origine (cfr. Rešetar 1997 [1911]). Nello SLM, però, l'ordine degli elementi della perifrasi è fisso: l'ausiliare precede sempre il verbo principale e può occupare anche la prima posizione nella frase, a differenza di quanto accade in croato standard, dove, come si è detto (par. 2.1), il clitico non può occupare la posizione frontale nella frase.

Anche in ragione del diverso ordine degli elementi della perifrasi, in SLM non si presenta nessuna forma agglutinata o fusa del clitico dell'ausiliare. Il futuro con *tit* in SLM, dunque, presenta un livello di

⁶ Tranne che per la prima persona singolare nella frase negativa, infatti, le forme piene del verbo *tit* sono utilizzate quando *tit* è impiegato con il suo significato proprio e non è in funzione ausiliare.

grammaticalizzazione meno avanzato rispetto al croato standard, in quanto in SLM non si è completato il processo di cliticizzazione relativo all'ausiliare e non è ancora avviato un processo di affissazione (Hopper e Traugott 1993: 108). Tuttavia, anche in SLM l'uso esclusivo della forma breve di *tit* nella perifrasi temporale (cfr. (8) con (9a) e (9b)) segnala la sua funzione ausiliare e evidenzia un certo grado di grammaticalizzazione di questa forma perifrastica.

- | | | | |
|-----|------------------------------|---------------|-----------------|
| (8) | <i>č-u</i> | <i>lej-it</i> | <i>džurnal'</i> |
| | volere.AUX-PRS.1SG | leggere-INF | giornale |
| | 'leggerò il giornale' | | |
| (9) | a. <i>hoč-em</i> | <i>lej-it</i> | <i>džurnal'</i> |
| | volere-PRS.1SG | leggere-INF | giornale |
| | 'voglio leggere il giornale' | | |
| | b. <i>hoč-em</i> | <i>na</i> | <i>džurnal'</i> |
| | volere-PRS.1SG | un | giornale |
| | 'voglio un giornale' | | |

Al contrario della forma con *tit*, il futuro con *imat* (7b) non ha riscontri nelle attuali varietà croate presenti nei Balcani⁸.

Anche nella perifrasi SLM con *imat* è presente la forma breve dell'ausiliare, che può anch'essa essere seguita dall'infinito di verbi sia perfettivi sia imperfettivi. A differenza di quanto accade con le forme brevi di *tit*, alla forma breve di *imat* può essere dato anche un

⁷ La trascrizione dei dati del corpus raccolti dagli informanti segue una convenzione ortografica fonetica, seppure attraverso la scelta di alcuni simboli della tradizione slava. Segnalo con *č* l'affricata palatale sorda; con *š* e *ž* rispettivamente la fricativa alveolare sorda e la sonora. Le convenzioni utilizzate si propongono di riportare il più fedelmente possibile le reali produzioni degli informanti. Le trascrizioni delle registrazioni, dunque, non seguono alcun criterio etimologico, discostandosi pertanto dalle convenzioni usate da Breu e Piccoli (2000). Per i dati presi da altri testi, invece, si è mantenuta l'ortografia usata dalla fonte.

⁸ L'uso dell'ausiliare *imat* (< *iměti*) nella formazione di strutture perifrastiche per il futuro, sebbene assente dal croato standard e in molte delle lingue dei Balcani, non è però assente nella storia delle lingue slave. Nell'antico slavo si realizzavano strutture perifrastiche con *iměti* 'avere' che insieme al valore temporale esprimevano significati modali di obbligo e necessità (Vaillant 1966: 106). Nel bulgaro-macedone, grazie all'influenza della struttura greco-romana corrispondente (con gr. *έχω* e lat. *habeo* 'avere'), la perifrasi verbale con *iměti* iniziò ad essere utilizzata con valore temporale di futuro e divenne la prima forma perifrastica di futuro nelle lingue slave dei Balcani, prima di essere sostituita dalla forma devolutiva.

significato pieno e non ausiliare. Infatti, se la forma non ridotta di *imat* significa ‘avere’, la sua forma breve ottenuta con caduta della vocale iniziale può essere utilizzata anche per veicolare il significato di ‘dovere’ oltre che per costruire la RTF in (7b). Questo dato evidenzia immediatamente un diverso livello di grammaticalizzazione della forma di futuro con *imat* rispetto a quella con *tir*⁹ e, insieme ad altri elementi, suggerisce di riconoscere la forma temporale con *imat* (che definiamo come *futuro debitivo*, vista la sua origine) esito del contatto dello SLM con le varietà romanze dell’area in cui è parlato¹⁰.

In merito alla forma (7c), si è visto che il presente indicativo con RTF è presente anche nel croato. Considerato anche il lungo e intenso contatto dello SLM con le varietà romanze meridionali, non meraviglia dunque la forma del presente indicativo tra le RTF dello SLM, già annotata con questa funzione anche da Rešetar (1997 [1911]: 136). Anche per il presente indicativo, la RTF può essere espressa sia da verbi imperfettivi sia da verbi perfettivi. Nelle produzioni dei parlanti abituali di SLM la forma del presente indicativo per designare eventi futuri è normalmente accompagnata da avverbiali che esplicitano il riferimento temporale (v. oltre); la mancanza di tale esplicitazione avverbiale permette quindi di assegnare una referenza temporale presente, eventualmente anche con valore aspettuale progressivo per i verbi imperfettivi.

4. *L’analisi sociolinguistica dei dati dello Slavo Molisano*

La descrizione delle forme di RTF dello SLM appena presentata si basa su dati raccolti sul campo in due dei tre comuni slavo-molisani, Acquaviva Collecroce e Montemitro (d’ora in poi rispettivamente AV

⁹ Situazione comune anche ad altre lingue. Si pensi, ad esempio, alle forme *shall* e *will* nella formazione del futuro inglese.

¹⁰ Gli antenati degli attuali croati molisani arrivarono in Italia tra fine XV e inizio XVI secolo partendo da aree croate in cui erano parlate varietà del gruppo dialettale štokavo (Rešetar 1997: 33) nelle quali fin dal XIII secolo la RTF era espressa con l’ausiliare ‘volere’ (Vaillant 1966: 109). Nel XIII secolo, infatti, la perifrasi volitiva si diffuse in gran parte dell’area balcanica, lasciando scomparire la precedente costruzione con ‘dovere’ che è limitata, oggi, solo al nord dell’Albania (Banfi 1985:153; 2003), in Bulgaria e nel rumeno balcanico (v. par. 2.1).

e MM), con diverse modalità di elicitazione¹¹. I dati di prima mano sono stati confrontati e integrati con quelli raccolti da Rešetar (1997 [1911]) a inizio Novecento e quelli presenti nei lavori di Breu (in particolare, 1992; 1993; 2003) e Breu e Piccoli (2000; 2011; 2012).

La maggior parte dei dati commentati in questo paragrafo nel quale propongo una analisi sociolinguistica dei dati raccolti, è stata elicitata attraverso un questionario traduttivo somministrato parzialmente una prima volta nel 2004 e poi nuovamente nel 2018 e nel 2019, che è servito anche da base per ulteriori produzioni e commenti anche di tipo metalinguistico da parte di alcuni informanti¹².

La raccolta è avvenuta in due comuni SLM che si differenziano notevolmente per diffusione della varietà slava e per atteggiamento nei confronti della varietà slava locale (v. anche Marra 2007, Bada 2009). Sia a MM sia ad AV la varietà slava è ancora diffusa e usata tra gli anziani e gli adulti, ma tra le generazioni più giovani si rileva una differenza di intensità d'uso differenziata nei due comuni. MM, infatti, ha una vitalità sociolinguistica (v. Unesco 2003) più elevata, con un uso dello SLM ancora diffuso anche tra i giovani (e tra questi anche delle ragazze) nei contesti informali e in alcuni contesti semi-formali intracomunitari; qui si riscontrano una trasmissione intergenerazionale abbastanza stabile, scelte di politica linguistica più attive, almeno fino a quando le risorse finanziarie lo hanno permesso, e un diffuso atteggiamento linguistico positivo e leale nei confronti dello SLM. AV, invece, presenta un minore uso dello SLM anche in contesti informali intracomunitari e un uso dell'italiano decisamente maggiore non solo tra i bambini ma anche tra i giovani, con una maggiore tendenza al passaggio alle varietà romanze da parte delle ragazze. Ad AV la trasmissione intergenerazionale è più compromessa e l'atteggiamento nei confronti dello SLM, sebbene migliorato nell'ultimo decennio grazie anche alle proposte di tutela e promozione della varietà slava locale, non appare del tutto positivo in parte della popolazione.

¹¹ La raccolta dati di prima mano, nella sua totalità, include traduzioni di brevi testi, descrizioni di vignette, conversazioni semi-guidate, conversazioni spontanee. Tutte le produzioni sono orali.

¹² Una parte dei dati e alcune delle osservazioni proposte in questo lavoro sono già in Marra (2005 e 2019). Tutti gli esempi presentati, se non è esplicitamente indicata altra fonte, sono stati raccolti di prima mano.

Date le differenze tra i due comuni slavofoni, nell'analisi dei dati da un punto di vista sociolinguistico si considereranno diversi parametri: comunità di appartenenza (AV o MM), alcune variabili demografiche (età e sesso dei parlanti) spesso correlate a processi di trasformazione linguistica a favore di varianti e varietà standard, distribuzione e frequenza d'uso dello SLM e delle altre varietà delle comunità (italiano e dialetto molisano) nel repertorio individuale degli informanti.

Gli informanti consultati sono 44. I dati, come si è già detto, sono stati elicitati attraverso diversi compiti. Non tutti gli informanti, però, hanno svolto tutti i compiti di elicitazione e non tutti hanno realizzato il test traduttivo messo a punto per verificare le forme di futuro trattate in questo contributo. Le mie riflessioni, pertanto, pur partendo da una analisi descrittivo-qualitativa delle forme di futuro riscontrate nell'intero corpus raccolto, per la verifica quantitativa si sono basate sulle produzioni di un sottocampione di informanti ai quali è stato chiesto di tradurre dall'italiano allo SLM venti brevi testi contenenti forme di futuro in contesti modali di vario tipo. Gli informanti del sottocampione sono stati selezionati in maniera tale da poter avere un continuum virtuale costruito sulla base dell'età e anche della presenza dello SLM nel loro repertorio che ha permesso di distinguere i parlanti abituali da quelli non-abituali¹³. Tra i primi sono stati inseriti coloro che hanno dichiarato e per i quali l'osservazione esterna ha verificato un uso abituale dello SLM (non necessariamente in forma esclusiva) nei vari contesti d'uso all'interno della comunità; tra i parlanti non abituali, al contrario, sono stati inseriti coloro che usano preferibilmente l'italiano in tutti i contesti¹⁴.

Per il task traduttivo sono stati registrati 28 informanti distribuiti inizialmente in cinque gruppi sulla base dei parametri appena evidenziati, ridotti poi a quattro (v. Tabella 1) vista l'uniformità di risultati dei parlanti SLM abituali anziani e di quelli adulti. Per MM, invece,

¹³ All'uso abituale si accompagna di solito anche una produzione che possiamo definire empiricamente fluente, al contrario di quanto accade con i parlanti non abituali che possono essere considerati in alcuni casi semi-parlanti (Dorian 1977).

¹⁴ La situazione linguistica familiare di questa tipologia di parlanti è varia: alcuni di loro hanno genitori slavofoni con i quali comunicano per lo più in italiano, alcuni hanno genitori che anche tra loro tendono ad usare l'italiano (e anche il dialetto molisano) ma non lo SLM.

il numero degli intervistati è molto minore, solo 8 (v. Tabella 1), ma la maggiore uniformità delle produzioni riscontrate in questo comune non evidenzia nessuna problematicità nel numero più ridotto di informanti.

Tabella 1 - *Numero di informanti che hanno svolto il task traduttivo sul futuro*

<i>Gruppi di parlanti</i>	<i>AV</i>	<i>MM</i>
parlanti abituali di SLM anziani (oltre i 60 anni) e adulti (26-59 anni)	10	4
parlanti abituali giovani (14-25 anni)	8	4
parlanti non abituali giovani (14-25 anni)	3	-
parlanti non abituali giovanissimi (meno di 14 anni)	7	-

4.1 La distribuzione delle forme di futuro in SLM

La raccolta e l'analisi dei dati parte dall'ipotesi che le due forme di futuro devolutivo e futuro debitivo non siano più usate solo in considerazione della diversa dimensione modale originaria (come sembra fossero un secolo fa; cfr. Rešetar 1997 [1911]), ma anche in relazione al ruolo che le varietà romanze giocano nel repertorio di ciascun informante e in quello della comunità a cui l'informante appartiene. In questa prospettiva si inserisce anche il rilevamento delle forme del presente indicativo, indicate anch'esse a inizio secolo scorso come espressione della RTF con la restrizione dell'esplicitazione dell'elemento avverbiale temporale, tenendo conto che sono ampiamente utilizzate con RTF anche nelle varietà romanze locali (v. par. 3).

Sulla base delle diverse variabili indipendenti considerate, l'analisi delle produzioni raccolte con il *task* traduttivo mostra che i parlanti abituali anziani e adulti di AV tendono a usare la forma devolutiva nei contesti in cui prevale il valore temporale (10)¹⁵ e la forma debitiva in contesti in cui permane un valore modale di obbligo o necessità (11), con una distribuzione funzionale delle due forme corrispondente a quanto riscontrato nei dati di inizio secolo scorso raccolti da Rešetar (1997 [1911]: 141).

¹⁵ Per ogni esempio del corpus è utilizzata una sigla che segnala il sesso del parlante (Donna, Uomo), l'età espressa in anni nel momento della raccolta dati, il comune di appartenenza (AV = Acquaviva Collecroce, MM = Montemitro) e se è un parlante abituale (Ab) o non abituale/semi-parlante (nonAb).

(10) Donna, 33_AV_Ab

<i>ne</i>	<i>č-eju</i>	<i>č-it</i>	<i>garu</i>
NEG	volere.AUX-PRS.3PL	fare-INF	gara
<i>si</i>	<i>dažd-i</i>		
se	piovere- PRS.3SG		
'non faranno la gara se piove'			

(11) Uomo, 71_AV_Ab

<i>ma-m</i>	<i>frunt-at</i>	<i>meštricu</i>	<i>do</i>
dovere-PRS.1SG	incontrare- INF	maestra	di
<i>moje</i>	<i>čere</i>	<i>sutr</i>	<i>istr</i>
mia	figlia	domani	mattina
'domani mattina incontrerò la maestra di mia figlia / domani mattina devo incontrare la maestra di mia figlia'			

I parlanti abituali più giovani, invece, ampliano l'uso della forma con *imat* anche a contesti non interpretabili come debitivi (12).

(12) Donna, 19_AV_Ab

<i>Sa</i>	<i>gled-am</i>	<i>televizion</i>	<i>pa</i>
ora	guardare-PRS.1SG	televisione	poi
<i>m-am</i>	<i>lej</i>	<i>na</i>	<i>džurnal</i>
dovere-PRS.1SG	leggere-INF	un	giornale
'Ora guardo la televisione, poi leggerò un giornale'			

Nei parlanti non abituali (giovani e giovanissimi) le forme perifrastiche devolutive sono ancora meno frequenti, favorendo una espansione d'uso della forma debitiva che, evidentemente, in questa tipologia di informanti ha ormai perso (o tende sempre più a perdere) il suo valore modale (13), soprattutto quando l'esposizione allo SLM è molto ridotta anche nel contesto familiare. Tra i parlanti non abituali, infatti, solo i giovanissimi che hanno genitori e nonni stabilmente slavofoni utilizzano ancora alcune forme devolutive.

(13) Uomo, 10_AV_non Ab

<i>m-am</i>	<i>po</i>	<i>u mor</i>
dovere-PRS.1SG	andare-INF	a mare
'andrò al mare'		

Per quanto riguarda la RTF espressa dal presente indicativo, tra i parlanti abituali di tutte le età non ci sono significative differenze: da tutti loro questa forma è usata per esprimere la RTF in concomitanza con avverbiali che esplicitano la deissi temporale (14), per lo più in

contesti in cui è implicata anche intenzionalità del soggetto o una dimensione di programmazione.

- (14) Donna, 35_AV_Ab
sutr vid-im Mariu
 domani vedere-PRS.1SG Maria
 ‘domani vedrò Maria’

Nei semi-parlanti, invece, si registra un aumento d’uso della forma del presente (v. Tabella 2) anche senza che ci sia nessuna esplicitazione temporale di accompagnamento (15).

- (15) Uomo, 10_AV_non Ab
jes-am dom
 essere-PRS.1SG casa
 ‘sarò a casa’

Tabella 2 - *Distribuzione delle forme di futuro nelle traduzioni degli informanti di AV*

F. di futuro	P. abituali anziani e adulti		P. abituali giovani		P. non-abituali giovani		P. non-abituali giovannissimi	
	n.	%	n.	%	n.	%	n.	%
<i>tit</i>	96	40	41	21	0	0	12	7
<i>imat</i>	92	38	104	54	45	61	102	60
presente	54	22	47	25	29	39	55	33
totale	242	100	192	100	74	100	169	100

I dati di MM sono stati per ora raccolti solo da parlanti abituali. Le forme finora attestate hanno evidenziato una situazione inattesa. Vista la maggiore diffusione e conservazione dello SLM tra gli abitanti di MM ipotizzavo per tutti gli informanti abituali una maggiore stabilità della distribuzione delle RTF rispetto ai dati di Rešetar di inizio secolo scorso, al pari di quanto riscontrato nei parlanti abituali anziani e adulti di AV. I dati di MM, invece, mostrano che la distribuzione d’uso delle tre forme si è ampiamente sbilanciata a favore della perifrasi volitiva (v. Tabella 3) che è utilizzata anche in contesti in cui è evidente la presenza di una modalità debitoria o nei quali sono espressi valori di intenzionalità e programmazione ed è presente una esplicitazione attraverso forme avverbiali (16).

- (16) Donna, 50_MM_Ab
di solito id-em u Termule
 di solito andare- PRS.1SG a Termoli
rab-it dana č-u po-kj Kambavaš
 lavorare-INF oggi volere.AUX-PRS.1SG andare-INF Campobasso
 ‘di solito vado a lavorare a Termoli; oggi andrò a Campobasso’

Questo sbilanciamento appare ancora più evidente nel confronto tra i parlanti abituali giovani dei due comuni (cfr. Tabelle 2 e 3).

Tabella 3 - *Distribuzione delle forme di futuro nelle traduzioni degli informanti di MM*

Forme di futuro	P. abituali anziani e adulti		P. abituali giovani	
	n.	%	n.	%
<i>tit</i>	79	82	83	85
<i>imat</i>	0	0	3	3
presente	17	18	12	12
totale	96	100	98	100

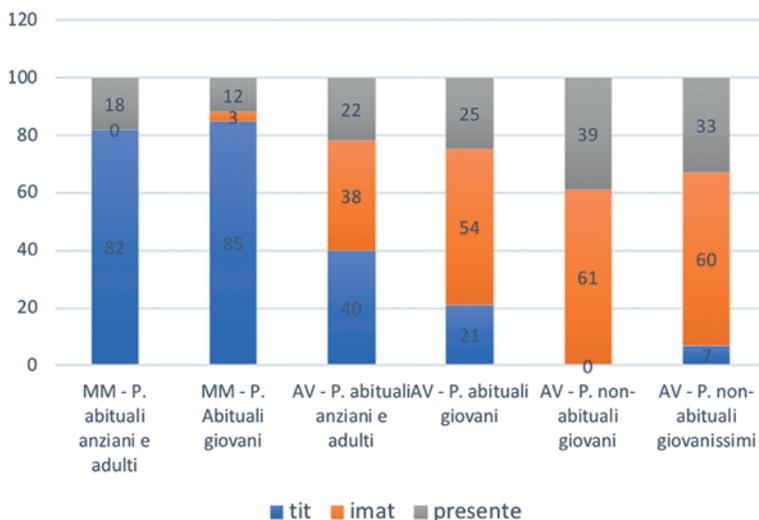
Sebbene sia opportuno una verifica di questa preferenza per la struttura devolutiva a MM anche attraverso produzioni più spontanee, i numeri attualmente riscontrati in alcuni dati spontanei da me già raccolti e quelli del corpus di Breu e Piccoli (2012) solo parzialmente analizzati sembrano confermare la tendenza. Anche in questi dati non conteggiati per questo contributo, infatti, le perifrasi devolutive sono ampiamente più diffuse e sono utilizzate anche insieme a forme avverbiali esplicite in contesti di programmazione e predittibilità che potrebbero prevedere l'uso del presente indicativo. Al contrario, le produzioni che presentano il modale 'dovere' (nella forma breve di *imat*) hanno un valore modale preponderante (anche se è presente, talvolta, una RTE, che però certamente non è prioritaria; v. ad esempio *maš kalat* in (17)). Infine, a MM sono pochissime e sempre accompagnate da esplicitazione temporale le occorrenze di presente.

- (17) Uomo,69_MM_Ab
 [...] *vičeras m-aš kal-at dol di masarija*
 [...] stasera dovere-PRS.2SG scendere-INF giù dove masseria
naša e mi pa č-mo do veče kasno
 nostra e noi poi volere.AUX-PRS.1PL andare-INF più tardi
 ‘[...] stasera devi scendere /scenderai giù alla nostra masseria. E
 poi noi verremo più tardi’ (Breu & Piccoli 2012: 20)

4. Conclusioni

Il confronto tra i dati di AV e MM (v. Grafico 1) e quelli raccolti a inizio Novecento permette di evidenziare una serie di modifiche e riorganizzazioni in atto nel sistema di RTF dello SLM.

Grafico 1 – Confronto in percentuale dei dati nei due comuni SLM



I mutamenti riscontrati mostrano che le trasformazioni del sistema possono essere sensibili a diversi fattori e possono attivare trasformazioni molto diverse anche all'interno di sottocomunità ristrette e tra loro non isolate.

Ad AV si è realizzata una evidente estensione delle funzioni della perifrasi debitoria con *imat* rispetto a inizio Novecento a discapito della RTF devolutiva con *tit*. Non è al momento possibile verificare se l'aumento del valore temporale della struttura debitoria sia causa o ef-

fetto di questa sovrestensione d'uso. I dati raccolti permettono solo di riscontrare che il valore modale di questa forma appare molto affievolito negli usi dei giovani parlanti abituali di AV rispetto a quanto lo sia nei parlanti abituali anziani e adulti della stessa comunità. Tuttavia, l'ulteriore restrizione d'uso della struttura devolutiva con *tit* (insieme alla estensione d'uso anche della RTF con presente indicativo) tra i semi-parlanti induce a ipotizzare che il contatto con le varietà romanze del territorio abbia favorito il processo di grammaticalizzazione della forma debitiva.

Se così fosse, le trasformazioni riscontrate tra i parlanti di MM potrebbero essere interpretate come esito di un processo di ristrutturazione del sistema delle RTF di natura opposta: una sorta di segnalazione di distanziamento dalle varietà romanze di contatto. Il mutamento riscontrato a MM produce un risultato che possiamo definire paradossale, in quanto la varietà di SLM più vitale sul piano sociolinguistico sta riorganizzando il sistema di RTF determinando una riduzione, di fatto, delle opzioni messe a disposizione dal sistema linguistico.

I dati di MM, dunque, appaiono decisamente più sorprendenti di quelli di AV. In quest'ultima comunità, infatti, il processo di trasformazione delle forme di futuro appare in linea con la dinamicità e ciclicità tipica di queste strutture riscontrate in diverse lingue del mondo (v. par. 2.1; cfr. tra gli altri Bybee & Pagliuca 1987; Bybee *et al.* 1994; e più in particolare Fleischman 1983 e Schwegler 1990 per le lingue romanze; Vaillant 1966, Banfi 1985 e Tomić 2004 per l'area balcanica). Al contrario, la direzione di trasformazione tracciata dalle produzioni dei parlanti di MM non risponde agli schemi più diffusi: in questo caso, infatti, il contatto interlinguistico non ha prodotto avvicinamento e convergenza dei sistemi ma ha determinato una reazione opposta, che tuttavia porta ad una semplificazione del sistema delle RTF. Un risultato, questo, non banale per le ricerche sul contatto linguistico, che invece evidenziano una tendenza alla maggiore complessità nelle varietà meno interferite (v. tra gli altri, Dorian 1981, Sasse 1992 e più recentemente, tra i molti, Trudgill 2009; Maitz & Németh 2014).

Ringraziamenti

Ringrazio molto i revisori anonimi per i loro preziosi commenti e suggerimenti. Solo mia, ovviamente, è la responsabilità di eventuali errori e lacune ancora presenti.

Riferimenti bibliografici

- Bada, Maria. 2009. La minoranza croata del Molise: un'indagine sociolinguistica e glottodidattica. In Franceschini, Rita (a cura di), *Le facce del plurilinguismo: fra metodologia, applicazione e neurolinguistica*, 100-169. Milano: Franco Angeli.
- Banfi, Emanuele. 1985. Isogrammatismo e calco linguistico: il cambio del futuro nelle lingue balcaniche. In Agostiniani, Luciano & Bellucci Maffei, Patrizia & Paoli, Matilde (a cura di), *Linguistica storica e cambiamento linguistico*, 147-159. Roma: Bulzoni.
- Banfi, Emanuele. 2003. Tipi di futuro in area Balcanica: questioni tipologiche, areali e storico-linguistiche. In Loi Corvetto, Ines (a cura di), *Dalla linguistica areale alla tipologia linguistica*, 165-206. Roma: Bulzoni.
- Bertinetto, Pier Marco. 1991. Il verbo. In Renzi, Lorenzi & Salvi, Giampaolo & Cardinaletti, Anna (a cura di), *Grande grammatica italiana di consultazione, Vol. II, I sintagmi verbale, aggettivale, avverbiale. La subordinazione*, 13-161. Bologna: Il Mulino.
- Breu, Walter. 1992. Das italo-kroatische Verbsystem zwischen slavischem Erbe und kontaktbedingter Entwicklung. In Reuther, Tilmann (a cura di), *Slavistische Beiträge Band 292, Slavistische Linguistik 1991: Referate des XV Konstanzer Slavistischen Arbeitstreffens*, 93-122. München: Otto Sagner.
- Breu, Walter. 1993. Verben der Fortbewegung im Italo-kroatischen in vergleichender Sicht (Morphologie, Funktionen, Entlehnungen, Rektion). In Kempgen, Sebastian (a cura di), *Slavistische Beiträge Band 304, Slavistische Linguistik 1992: Referate des XVIII Konstanzer Slavistischen Arbeitstreffens*, 9-40. München: Otto Sagner.
- Breu, Walter. 2003. Bilingualism and linguistic interference in Slavic-Romance contact area of Molise (Southern Italy). In Eckardt, Regine & von Heusinger, Klaus & Schwarze, Christophe (a cura di), *Words in time. Diachronic semantics from different points of view*, 351-373. Berlin-New York: Mouton de Gruyter.

- Breu, Walter & Piccoli, Giovanni. 2000. *Dizionario Croato Molisano di Acquaviva Collecroce. Dizionario plurilingue della lingua slava della minoranza di provenienza dalmata di Acquaviva Collecroce in provincia di Campobasso. Dizionario, registri, grammatica, testi*. Ripalimosani (Cb): Arti Grafiche La Regione.
- Breu, Walter & Piccoli, Giovanni. 2011. *Südslavisch unter romanischem Dach. Teil I. Texte aus Acquaviva Collecroce*. München-Berlin-Washington (D.C.): Otto Sagner.
- Breu, Walter & Piccoli, Giovanni. 2012. *Südslavisch unter romanischem Dach. Teil II. Texte aus Montemitro und San Felice del Molise*. München-Berlin-Washington (D.C.): Otto Sagner.
- Browne, Wayles. 1993. Serbo-Croat. In Comrie, Bernard & Corbett, Greville G. (a cura di), *The Slavonic languages*, 305-387. London-New York: Routledge.
- Bybee, Joan L. & Pagliuca, William. 1987. The evolution of future meaning. In Giacalone Ramat, Anna & Carruba, Onofrio & Bernini, Giuliano (a cura di), *Papers from the 7th international conference on historical linguistics*, 109-122. Amsterdam-Philadelphia: John Benjamins.
- Bybee, Joan L. & Perkins, Revere & Pagliuca, William. 1994. *The evolution of grammar. Tense, aspect, and modality in the languages of the world*. Chicago-London: The University of Chicago Press.
- Chung, Sandra & Timberlake, Alan. 1985. Tense, aspect and mood. In Shopen, Timothy (a cura di), *Language typology and syntactic description, Vol. III. Grammatical categories and the lexicon*, 202-258. Cambridge-London-New York: Cambridge University Press.
- Comrie, Bernard. 1985. *Tense*. Cambridge-London-New York: Cambridge University Press.
- Comrie, Bernard & Corbett, Greville G. (a cura di). 1993. *The Slavonic languages*. London-New York: Routledge.
- Dahl, Östen. 2000. The grammar of future time reference in European languages. In Dahl, Östen (a cura di), *Tense and aspect in the languages of Europe. Empirical approaches to language typology*, EURO TYP 20-6, 309-328. Berlin-New York: Mouton de Gruyter.
- De Roberto, Elisa. 2010. Futuro. In *Enciclopedia dell'Italiano Treccani*, <

- Dorian, Nancy C., 1981. *Language death: the life cycle of a Scottish Gaelic dialect*. Philadelphia: University of Pennsylvania Press.
- Fleischman, Suzanne. 1983. From pragmatics to grammar. Diachronic reflections on complex pasts and futures in Romance. *Lingua* 60. 183-214.
- Hopper, Paul J. & Traugott, Elizabeth C. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- Kaye, Steven. 2009. *Grammaticalization and future reference in the Italic/Romance verb*. Oxford: University of Oxford (Tesi di Master of Philosophy).
- Leiss, Elisabeth & Abraham Werner (a cura di). 2014. *Modes of modality. Modality, typology and Universal grammar*. Amsterdam-Philadelphia: John Benjamins Publishing Company.
- Lokmane, Ilze & Kalnača, Andra. 2014. Modal semantics and morphosyntax of the Latvian DEBITIVE. In Leiss, Elisabeth & Abraham, Werner (a cura di), 167-192.
- Maitz, Péter & Németh, Attila. 2014. Language contact and morphosyntactic complexity: evidence from German. *Journal of Germanic Linguistics* 26(1). 1-29.
- Marra, Antonietta. 2005. Mutamenti e persistenze nelle forme di futuro dello slavo molisano. In Breu, Walter (a cura di), *L'influsso dell'italiano sulla grammatica delle lingue minoritarie. Problemi di morfologia e sintassi. Atti del Convegno internazionale, Costanza, 9-11 ottobre 2003*, 141-166. Rende: Università della Calabria - Centro Editoriale e librario.
- Marra, Antonietta. 2007. Politiche linguistiche e piccole comunità minoritarie, tra sociolinguistica e glottodidattica. In Consani, Carlo & Desideri, Paola (a cura di), *Minoranze linguistiche. Prospettive, strumenti, territori*, 161-185. Roma: Carocci.
- Marra, Antonietta. 2019. La riorganizzazione del sistema verbale nello Slavo del Molise. Su alcune forme di futuro. *Rhesis. International Journal of Linguistics, Philology, and Literature* 10(1). 56-74.
- Mauri, Caterina & Sansò, Andrea. 2012. What do languages encode when they encode reality status? *Language Sciences* 34. 99-106.
- Migdalski, Krzysztof Marek. 2006. *The syntax of compound tenses in Slavic*. Utrecht: LOT.
- Mitkovska, Liljana & Bužarovska, Eleni. 2014. Deontic or epistemic? *Habère* as a modal marker of future certainty in Macedonian. In Leiss, Elisabeth & Abraham, Werner (a cura di), 193-218.

- Nuyts, Jan. 2016. Analyses of the modal meanings. In Nuyts, Jan & Van der Awera, Johan (a cura di), *The Oxford handbook of modality and mood*, 31-49. Oxford: Oxford University Press.
- Palmer, Frank R. 2001. *Mood and modality*. Cambridge-New York-Melbourne-Madrid: Cambridge University Press.
- Rešetar, Milan. 1997 [1911]. *Le Colonie Serbo-Croate nell'Italia Meridionale* [Die Serbokroatischen Kolonien Südtaliens], Breu, Walter & Gardenghi, Monica (a cura di). Campobasso: Amministrazione Provinciale di Campobasso.
- Rohlf, Gerhard. 1968 [1949]. *Grammatica storica della lingua italiana e dei suoi dialetti. Morfologia*. Torino: Einaudi.
- Sansò, Andrea. 2020. Routes towards the irrealis. *Transactions of the Philological Society* 118(3). 401-446.
- Sasse, Hans-Jürgen. 1992. Language decay and contact induced change: similarities and differences. In Brenzinger, Matthias (a cura di), *Language death. Factual and theoretical explorations with special reference to East Africa*, 59-80. Berlin-New York: Mouton De Gruyter.
- Schwegler, Armin. 1990. *Analyticity and syntheticity. A diachronic perspective with special reference to Romance languages*. Berlin-New York: Mouton De Gruyter.
- Tomić, Olga Mišeska. 2004. The syntax of the Balkan Slavic future tenses. *Lingua* 114(4). 517-542.
- Trudgill, Peter. 2009. Sociolinguistic typology and complexification. In Sampson, Geoffrey & Gil, David & Trudgill, Peter (a cura di), *Language complexity as an evolving variable*, 98-109. Oxford: Oxford University Press.
- UNESCO ad hoc expert group on endangered languages. 2003. *Language vitality and endangerment*. UNESCO, <<https://unesdoc.unesco.org/ark:/48223/pf0000183699>>.
- Vaillant, André. 1966. *Grammaire comparée des langues slaves. III. Le Verbe*. Paris: Klincksieck.

EMANUELE MIOLA

Varietà marginali di varietà substandard: appunti tipologici sui gerghi gallo-italici

Seguendo l'approccio proprio della tipologia sociolinguistica, il contributo prende in esame qualche tratto morfosintattico caratteristico dei gerghi gallo-italici e istituisce un confronto con i tratti corrispondenti del continuum dialettale ospite. Sulla base di un campione di dati desunti da vari lavori che si sono occupati delle varietà gergali, si osserva che, nel microsistema dei pronomi personali, nella negazione di frasi e nella struttura di sostantivi e aggettivi, i gerghi non esibiscono gli attesi mutamenti complicanti rispetto ai microsistemi delle varietà ospiti, anzi vi sono apprezzabili fenomeni di semplificazione. Si conclude quindi con una proposta di revisione e messa a punto della cosiddetta "congettura di Trudgill".

Parole chiave: tipologia sociolinguistica, complessificazione linguistica, gerghi, gallo-italico.

1. *Introduzione: tipologia sociolinguistica e gerghi*

Scopo di questi appunti è analizzare qualche microsistema grammaticale di alcuni gerghi delle varietà gallo-italiche e valutarne la maggiore o minore complessità rispetto ai corrispondenti microsistemi del continuum dialettale ospite seguendo l'approccio sociolinguistico-tipologico proposto in molti lavori da Trudgill (per es. 2009, 2011). In altre parole, in questo contributo si proverà a saggiare le ipotesi di Trudgill relative ai fattori sociali che giocano un ruolo nel rendere una varietà linguistica più complessa o, al contrario, più semplice di un'altra seguendo l'acuto suggerimento di Moretti (2017: 40-41), il quale, proprio nell'ambito di una discussione critica delle proposte di Trudgill, rilevava che poteva diventare «estremamente interessante [...] l'indagine di forme di complessità spontanea che possono apparire in sottogruppi appartenenti sì alla macro-società, ma caratterizzati proprio da caratteristiche di forte chiusura, numero di membri ridotto, ecc., [...] come nel caso dei gerghi».

Come è ben noto, i gerghi, nell'architettura della lingua che li ospita, risultano essere varietà marcate come meno prestigiose sia in diastratia sia in diafasia, in quanto sono tipiche di categorie o ristretti gruppi di utenti variamente identificabili come subalterni e socialmente marginali e sono inoltre caratterizzate da massima informalità e da un uso *in-group* quasi esclusivo (Berruto [1987: 25>]2012: 28; Sanga 1993: 153-155; Marcato 2013: 10-13; Scala 2016: 47-50). Per definizione, dunque, il gergo è parlato da un gruppo numericamente non elevato, che costituisce una parte minoritaria della comunità che parla la varietà ospite. I gerganti generalmente sono legati da legami sociali forti, hanno un'alta quota di conoscenze condivise e, in quanto «glottologi spontanei», per dirla con Sanga (1984: 12), sono in contatto con altre lingue e altre varietà che, per di più, spesse volte imparano e conoscono. Esamineremo in particolare alcuni gerghi storici, cioè quelli praticati da mendicanti, malviventi, vagabondi, ambulanti, e alcuni gerghi meccanici a modificazione sillabica, tutti ospitati dal cosiddetto 'diasistema' o 'gruppo dialettale' gallo-italico (Bossong 2016: 68), che include le varietà piemontesi, lombarde, della Svizzera italiana ed emiliano-romagnole¹. Anche se le varietà indagate sono limitate a un particolare territorio, sono tuttavia convinto che i dati discussi restituiranno un'immagine che si possa attagliare ai gerghi in generale, in ragione della loro sostanziale unitarietà (v. tra i molti altri Sanga 1984: 189-195).

Gli appunti che seguono si articolano in tre sezioni. Dapprima verranno offerte alcune definizioni di riferimento per intendere il discorso che si andrà dipanando: a che cosa ci riferiamo – almeno per lo spazio di questo scritto – con complessificazione e semplificazione linguistica e con le etichette 'più complesso' e 'più semplice'. Si mostreranno quindi i dati di tre microsistemi gergali e li si comparerà con quelli delle varietà gallo-italiche loro ospiti². Nell'ultimo paragrafo – se è concesso anticipare già qui le conclusioni – si prenderà atto che

¹ Pur se nella manualistica il ligure viene talvolta ascritto al gruppo gallo-italico, le sue caratteristiche linguistiche spingerebbero a escluderlo, v. Toso (2002).

² Un revisore anonimo si chiede in che misura sia lecito testare gli assunti di Trudgill sui gerghi, i cui utenti si dedicano soprattutto alla creazione consapevole di lessico che risulti esclusivo, senza investire quasi nulla su una grammatica esclusiva. Anche se è in effetti vero che i gerganti lavorano più sul materiale lessicale che sulle regole morfosintattiche delle varietà che conoscono, a riprova di quanto il sistema lingua sia fatto di sottodomini interconnessi, non sono da escludersi – e anzi come si vedrà si

i dati a nostra disposizione non mostrano chiari fatti di complessificazione tra le strutture del gergo e quelle delle varietà ospiti, com'era invece atteso sulla base degli studi di tipologia sociolinguistica, e si proverà a dare ragione di questo in ottica socio-tipologica suggerendo una possibile revisione delle ipotesi e delle proposte di Trudgill (2009, 2011, 2015).

2. *Complessità e semplificazione*

Prima di tutto occorre riconoscere che l'analisi condotta qui riposa su un particolare modo di intendere la complessità linguistica. Su questo tema, infatti, negli ultimi decenni la riflessione degli studiosi si è concentrata tanto da produrre molti e differenti approcci al problema, indicando pertanto anche differenti soluzioni. Per motivi di spazio non posso passare in rassegna tutti questi approcci e – senza alcuna pretesa di esaustività – potrò solo citare *en passant* l'articolo di McWhorter (2001), che ha riacceso l'interesse per la possibile eguale complessità di tutte le lingue, e i contributi di Kusters (2003), Dahl (2004), Miestamo *et al.* (2008), Sampson *et al.* (2009) e Mufwene *et al.* (2017), tra i molti altri che meriterebbero di essere ricordati. Un'ottima e approfondita panoramica sul tema è offerta da Fiorentino (2019: specie 37-45)³.

Scelgo qui di trattare la complessità, come già accennato, così come è intesa da Peter Trudgill, anche se ogni tanto potrò fare qualche cenno ad altri approcci. Sintetizzando molto, per Trudgill (2010: 307-308; 2011: 62-63) la complessità o meglio la complessificazione si verifica nelle lingue all'aumentare di irregolarità e di opacità, specialmente morfosintattica, all'aumentare di ridondanza e di ripetizione dell'informazione e all'aumentare delle categorie morfologiche.

È importante anche menzionare, almeno per i fini che mi prefiggo, i concetti di complessità spontanea e di complessità di seconda fase o secondaria (v. Moretti 2017). Questo secondo tipo di complessità è quello che scaturisce dall'introduzione della scrittura all'interno di una comunità e dalla conseguente maggiore stratificazione e com-

danno – anche risvolti morfosintattici a mutamenti che interessano in prima battuta il lessico.

³ Della stessa autrice si veda anche Fiorentino (2009), con un'elegante tabella riassuntiva a p. 284.

plessità delle società. Indagare le varietà gergali gallo-italiche ci permette perciò di ridurre al minimo le conseguenze della complessità secondaria, dal momento che guardiamo a sistemi linguistici che hanno sempre goduto di un basso o al massimo medio-basso grado standardizzazione scritta. Per quanto riguarda la stratificazione sociale, essa è praticamente ineludibile, giacché una qualche forma di questa è comunque presente in ogni comunità e in ogni società. Osservando le varietà qui in esame sarà quindi possibile concentrarci esclusivamente sulla complessità spontanea dei sistemi che chiamiamo lingue. Sarà dunque questa a suo modo un'indagine sulla complessità così come si è sempre presentata nella maggior parte dei casi nelle lingue del mondo almeno 'fino a ieri', per citare Diamond (2012), dato che in passato le lingue dotate di standardizzazione o anche solo di grafizzazione erano in numero molto ridotto rispetto a quelle esclusivamente parlate, anche se questo panorama, con il consolidarsi della comunicazione mediata dalla Rete, sembra essere in evoluzione (tra i molti altri, v. Patrucco 2003, Fiorentino 2005, Wright 2006, ma si tengano in conto anche i dubbi sollevati successivamente da Kornai 2013). Come che sia, per quegli oggetti linguistici che definiamo lingue, almeno dal mero punto di vista statistico, quelle di cui si discute qui sono state e sono ancora oggi le condizioni di complessificazione o semplificazione normali.

In questo quadro i gerghi sono interessanti perché le comunità che li parlano, o forse meglio che li parlavano, appartengono, come già detto al § 1, a quelle *societies of intimates* che esibiscono tutti o quasi tutti i tratti sociali collegabili ai fattori che secondo la cosiddetta "congettura di Trudgill" (2012: 92) sono responsabili della complessificazione delle lingue: «small size, dense social networks, large amounts of shared information, high stability, and low contact». Si badi, tutti o quasi tutti i tratti sociali: in effetti, i gerganti, essendo girovaghi, nomadi, lavoratori ambulanti o stagionali, hanno come caratteristica sociologica definitoria l'instabilità (Sanga 1993: 155-156), che porta necessariamente a frequenti occasioni di contatto con altre lingue e varietà linguistiche. Nonostante ciò, secondo la "congettura di Trudgill" ci sarebbe da attendersi comunque almeno un moderato grado di complessificazione dei gerghi rispetto alle lingue ospiti.

I dati che esporrò nel prossimo paragrafo serviranno a verificare e a discutere questa ipotesi.

2.1 L'autenticità dei dati gergali

È ancora opportuno, prima di entrare *in medias res*, spendere qualche parola sull'autenticità dei dati che stiamo per scrutinare.

Riprendendo le parole di Menarini (1942: 155), «il vero testo del gergo è la frase». Se da una parte questo ammonisce a non limitarsi a interrogare i gerganti chiedendo loro singole parole, dall'altra fa emergere come non sia facile elicitarle di queste varietà testi lunghi senza diminuirne almeno in parte l'autenticità. In effetti, «[s]olo dalla fine dell'Ottocento, con lo sviluppo delle ricerche linguistiche e antropologiche, possiamo avere una documentazione autentica e sicura del gergo» (Sanga 1993: 173), ma i «materiali per lo studio della sintassi, della testualità e della pragmatica gergale si devono cercare in fonti letterarie, preziose ma anche infide» (Sanga 1993: 171). Anche le poche, interessantissime, lettere gergali di cui disponiamo sollevano il dubbio di autenticità: *in primis*, non sono lingua parlata, come vorrebbe Menarini; poi, non è difficile obiettare che possono essere state scritte con il fine di accentuare la componente criptica del gergo, per essere sicuri quanto più possibile di non incorrere nella censura, come nel caso delle due *Kriegsgefangenenbriefen* in tarón che si leggono in Spitzer (2014: 30-31), oppure per ciò che potremmo chiamare esigenza dell'arcaismo, del preziosismo e del distanziamento rispetto alla varietà ospite, che chi conosce un gergo può mettere in atto anche solo perché immagina che il ricercatore che gli chiede, poniamo, una traduzione della parabola del figliol prodigo se li aspetti (si v. Iannàccaro 2000: specie 61-65). Infine, benché il gergo sia «usato [...] quasi esclusivamente per la comunicazione normale, corrente, tra gerganti» (Sanga 1993: 153), è raro avere registrazioni di gerganti che parlano tra loro senza che vi sia almeno un gagio astante, cosa che in un modo o nell'altro influisce sulla lingua adoperata nello scambio.

Insomma, quando si vuole osservare il gergo dal punto di vista linguistico, occorre tenere a mente che non abbiamo un vero e proprio corpus gergale interrogabile, che sono rare le descrizioni grammaticali gergali scientificamente accurate, che disponiamo di pochi – e in parte problematici – testi gergali scritti e di pochi dialoghi autentici tra gerganti e che da gergante a gergante, anche dello stesso gergo, può riscontrarsi una grande variazione idiolettale. In quanto segue si cercherà perciò di fare il miglior uso possibile di possibili *bad data*.

3. *La tipologia dei gerghi: tre sondaggi*

È noto che i gerghi italiani ed europei adoperino un lessico sostanzialmente unitario, pur se appoggiato alle caratteristiche fonetiche e grammaticali delle singole varietà ospite. Di maggiore interesse per noi sono, però, le caratteristiche morfosintattiche che si ritrovano in tutti i gerghi, che Sanga ha discusso in molti lavori e specialmente in Sanga (1993: 161ss). Prenderemo in considerazione tre di questi tratti morfosintattici unitari, rapportandoli poi ai corrispondenti sistemi della lingua ospite: il microsistema dei pronomi personali tonici, la costruzione della negazione di frasi e la struttura interna e dei paradigmi di nomi e aggettivi, occupandoci per quest'ultimo tratto soprattutto delle strategie derivazionali e dei meccanismi di permutazione e inversione sillabica che interessano alcune classi di parole gergali.

3.1 Il microsistema dei pronomi tonici

I gerghi tendono a esprimere il pronome personale tonico con una costruzione che può descriversi in questo modo: [DEF.ART + POSS + proforma]. L'articolo definito, naturalmente, si ha solo in quelle varietà che lo esprimono davanti al possessivo adnominale. Il possessivo riflette semanticamente la persona e il numero dell'intera forma pronominale, secondo il procedimento chiamato da Harris (1981: 51-52) e Collins *et al.* (2008) camuffamento (in inglese, *camouflage*). La proforma (o *mask*, maschera), infine, sarà un nome della varietà ospite, lessicalmente vuoto e che originariamente poteva fungere da nome proprio, oppure da nome comune, non di rado con il significato di 'corpo' o di una sua parte. A mo' di esempio, si veda (1), tratto da una varietà vernacolare di inglese:

- (1) African-American Vernacular English
his ass is gonna get fir-ed
 POSS.MSG culo COP.3SG FUT.AUX PASS.AUX.INF licenziare-PP
 'Verrà licenziato' (Spears 1998: 235)

Al posto del pronome personale di terza persona in (1) si adopera *his ass*, letteralmente '(il) suo culo'. La forma pronominale dell'AAVE è dunque in origine un sintagma nominale nello standard e assume i tratti di persona, genere e numero di *his*, mentre *ass* funziona come una sorta di metonimia per l'intera persona cui ci si riferisce.

Qualcosa di simile si incontra nel *rügín* dei calderai della Val Colla⁴, che presenta un sistema pronominale tonico a cinque unità. I pronomi diversi nel sistema ospite valcollese, che non «si discost[a] da quel tipo di lombardo-comune che s'ode a Lugano e nella sua campagna» (Salvioni 1891: 98), sono invece sette (dati dalle carte 660, 1253, 1607, 1627, 1660 dell' AIS, p. 73; Keller 1933, 1934: *passim*, 1937: 158; Papanti 1875: 629-630; da qui in avanti, in tutte le tabelle e gli esempi si adopera una trascrizione larga).

Tabella 1 - *Pronomi personali del rügín e del valcollese*

	<i>rügín</i>	<i>valcollese</i>
'io'	<i>el mè òden</i> ⁵	<i>mi</i>
'tu'	<i>el tò òden</i>	<i>ti</i>
'egli'	<i>el sò òden</i>	<i>lù</i>
'ella'	<i>el sò òden</i>	<i>lee</i>
'noi'	<i>el nòz òden</i>	<i>nün</i>
'voi'	<i>el vòz òden</i>	<i>vü</i>
'essi'/'esse'	<i>el sò òden</i>	<i>lur</i>

Condizioni perfettamente uguali, con cinque pronomi personali contro sette, si riscontrano comparando il tarüsc di Massino Visconti con il dialetto di quel paese o di quelli limitrofi, come Nonio (dati da AIS, p. 128 e Manni 1973: 13)⁶.

⁴ Oltre che in Keller, testi, glossari e studi di e sul *rügín* si leggono, tra gli altri, in Soldati (1957), Sanga (1977), Bertolotti *et al.* (1978), Lurati (1990), Arigoni/Vicari (2019) e Arigoni (2020).

⁵ *Òden* vale 'uomo, persona' per Sanga (1984: 213), sulla scorta anche di Cherubini (1843: s.v. *zerga*, p. 547). Si vedano anche Cortelazzo (1979) e Wagner (1990).

⁶ Tralascio qui di discutere gerghi con varietà ospiti dotate di sistemi pronominali ancora più numerosi, come quelle piemontesi, per le quali si potrebbe parlare di una flessione di genere anche per la prima e la seconda plurale, che più frequentemente che no presentano l'alternanza del tipo di *nojaut(r)i/nojaut(r)e* e *vojaut(r)i/vojaut(r)e*. Devo ad Andrea Scala la segnalazione dell'interessante caso dei pronomi personali del tarón della Val Rendena: la proforma qui è *gian*, che può diventare *giani* al plurale (Vigolo/Barbierato 2008). (*E*)*l mé gian* ecc e (*e*)*i nos(i) gian(i)* ecc. flettono differentemente dai nomi comuni del rendenese, dal momento che possono prendere l'uscita -i di alcuni elementi grammaticali come, appunto, i possessivi (v. per es. *nos-nosi*), forse per via dell'influsso della varietà di tarón e dei dialetti veneti del Trentino occidentale (Tomasini 1989: 32-33, 36-37).

Tabella 2 - *Pronomi personali del tarùsc e del noniese*

	<i>tarùsc</i>	<i>noniese</i>
'io'	<i>el me tona</i> ⁷	<i>mi</i>
'tu'	<i>el teu tona</i>	<i>ti</i>
'egli'	<i>el seu tona</i>	<i>lùii</i>
'ella'	<i>el seu tona</i>	<i>[lé]</i>
'noi'	<i>el neust tona</i>	<i>nüii</i>
'voi'	<i>el veust tona</i>	<i>vüii</i>
'essi'/'esse'	<i>el seu tona</i>	<i>lur</i>

I sistemi pronominali gergali, dunque, presentano un minor numero di unità, ma, a ben vedere, maggior materiale morfologico. In ogni *slot* del paradigma, a fronte di un singolo morfema della varietà ospite, i morfemi sono due: una base comune non continuativa (*el... òden* per fare il caso del *rügín*) con un infisso (il possessivo della varietà ospite) che indica la persona e, almeno per le prime due persone, il numero, ma non il genere. Questo perché in molti dialetti gallo-italici, come quelli visti nelle Tabelle 1 e 2, il possessivo concorda per genere non con il possessore, ma con il posseduto. Ne viene, almeno per i gerghi che abbiamo visto, un sistema pronominale che può essere, rispetto a quello delle varietà ospiti, ridotto anche sensibilmente, con neutralizzazione, inoltre, di ogni distinzione di numero alle terze persone. Per conseguenza si osserva anche una tendenziale riduzione del paradigma verbale, del quale si usa solo (o, meglio, molto più spesso) la terza persona singolare⁸.

Dal punto di vista tipologico, una costruzione di pronomi personali del tipo [base comune + affisso che marca la persona] è piuttosto rara nelle lingue del mondo, mentre molto più rappresentato è il tipo della varietà ospite, cioè quello con ciascuna cella occupata da un'unità diversa (Daniel 2013). Di contro, però, l'assenza di ogni distinzione di genere, che si verifica nei gerghi ma non nelle loro varietà ospiti, è nettamente il tipo più frequente nelle lingue del mondo (Siewierska 2013).

⁷ Dal nome proprio Antonio.

⁸ Qualche dato in merito: nella versione della parabola del figliol prodigo che si legge in Keller (1934) un terzo delle forme verbali che si riferiscono a una prima o una seconda persona singolare o plurale presenta l'uscita della terza persona (la terza plurale è foneticamente identica alla terza singolare anche nella varietà ospite).

In ogni caso, il decremento di irregolarità e la diminuzione del numero dei pronomi personali, e conseguentemente delle forme del paradigma verbale, sono indici di un processo di semplificazione, secondo i parametri di Trudgill (2015: 138ss), ma anche tra gli altri per Berruto (1990: 31), che parla appunto di «semplificazione paradigmatica in termini di ridotto numero delle unità presenti nell'inventario»⁹.

3.2. La frase negativa

Quanto alla frase negativa, si dirà subito che le varietà gallo-italiche ospiti possono essere a negazione solo postverbale, come è solito per il piemontese e il lombardo, oppure discontinua, come alcuni dialetti dell'emiliano-romagnolo (ess. 2-4).

- (2) Emiliano-romagnolo (bolognese)

a n al sò brisa
 SBJ.CL1SG NEG DO.CL3MSG sapere.PRS1SG NEG
 'Non lo so' (Vitali 2020: II 42)

- (3) Lombardo (gromese)

sta fomna la m pyaz mia
 questa donna SBJ.CL3FSG IO.CL1SG piacere.PRS3SG NEG
 stim.: 'questa donna non mi piace' (AIS 1678, p. 237)

- (4) Piemontese (torinese)

i capisso nen
 SBJ.CL1SG capire.PRS1SG NEG
 'non capisco' (Rohlf's 1969: 305, con modifiche)

Tratto caratteristico dei gerghi che stiamo indagando sarebbe la «costruzione delle frasi negative mediante il tipo *büs* "non"» (Sanga 1984: 192; 1993)¹⁰, oppure mediante altre parole usate dai gerganti per significare l'avverbio 'no', oppure una piccola quantità o entità, come il *rügín vita*, o ancora *löt*, del gianglamènt torinese (Aly Belfàdel 1898: 634), di etimologia non chiara. Queste parole sostituiscono il

⁹ Va aggiunto, tuttavia, che stiamo ragionando su una condizione per dir così ideale del gergo, in quanto quando si va sui (pochi, come detto) testi a disposizione, la grammaticalizzazione del 'nuovo' sistema pronominale costituito da [possessivo + proforma] e la generalizzazione delle forme verbali di terza persona singolare non risultano completamente compiute.

¹⁰ *Büs*, *boss* o *bus(chia)* valgono con tutta probabilità 'bastoncino', caso evidente di minimizzatori grammaticalizzati come elementi negativi.

morfema negativo postverbale della varietà ospite senza cambiare altro. Anzi, nel gergo i due morfemi negativi *vita* e *miga* convivono in un regime di competizione, come si vede dagli esempi (5-7):

- (5) Rügín
al voreva vita binà in kròspa
 SBJ.CL3MSG volere.IMPF3SG NEG andare.INF in casa
 ‘non voleva entrare in casa’ (Keller 1934: 77)
- (6) Rügín
Olpéga kwel ber léra, k al
 guardare.IMP2SG quello uomo là che SBJ.CL3MSG
ta faga miga binà i brünas
 IO.CL2SG fare.SBJV3SG NEG andare.INF DEF.ART ferri
 ‘fa’ attenzione che quel tizio non ti rubi i ferri del mestiere’
 (Keller 1934: 63)
- (7) Valcollese
a doerme miga
 SBJ.CL3MSG dormire.PRS3SG NEG
 ‘non dorme’ (AIS c. 653, p. 73)

Nel ladresco bolognese, il gergale *nekka* può essere usato al posto del morfema negativo postverbale del bolognese comune *brisa* (si confronti (8a) con (2)), che resta comunque il morfema negativo più frequente anche nel gergo (v. (8b)). *Nekka* può però comparire al posto di *bri(s)a* anche quando si trova come unico negatore in posizione preverbale, nell’imperativo negativo (v. (9a-b)).¹¹

- (8) a. Ladresco bolognese
a-n ñ e nekka la pilla
 SBJ.CL3SG-NEG LOC.CL è NEG DEF.ART denaro
 ‘il denaro non c’è’ (Menarini 1942: 100)
- (8) b. Ladresco bolognese
a-n n ø bris’a vòia
 SBJ.CL3SG-NEG PART.CL ho NEG voglia
 ‘non ne ho voglia’ (Menarini 1942: 40)

¹¹ *Nekka* nel gergo ha anche il significato di ‘niente’, in competizione con *ñinta* e con l’altro termine gergale *boss*. Per *nek(k)a*, parafrasi di ‘no’ e ‘niente’, v. anche Prati (1978: n. 251), con bibliografia, e Ferrero (1991: s.v. *nisba*). La voce è presente, come rafforzamento di frase negativa, anche nel vocabolario siciliano-italiano di Spatafora (cit. in Traina 1868: s.v. *neca*).

- (9) a. Ladresco bolognese
nekka bakaiɽ!
 NEG parlare.INF
 ‘non parlare!/stai zitto!’ (Menarini 1942: 42)
- (9) b. Ladresco bolognese
bria dɽ z'a al mi kulour
 NEG dare.INF giù DEF.ART POSS.1SG colore
 ‘non fare il mio nome!’ (Menarini 1942: 64)

Insomma, anche se i gerghi possono usare morfemi negativi non comuni e non frequenti nella varietà che li ospita, per quel che riguarda la costruzione della frase non si osservano differenze tipologiche: come i gerghi che parassitano varietà a sola negazione postverbale mantengono la stessa strategia, così – cosa più interessante per i fini della nostra indagine – i gerghi che parassitano varietà a negazione discontinua preservano la medesima strategia, anche se questa costruzione è ridondante e interlinguisticamente più rara (Dryer 2013)¹². In ogni caso, dal punto di vista della tipologia sociolinguistica, non si osserva né perdita di ridondanza né aumento di ridondanza. Forse si può dire, allora, che, per quanto riguarda questo microsistema grammaticale, i gerghi gallo-italici risultano meno trasparenti solo lessicalmente, e solo per i parlanti della varietà ospite. Questo perché, in fondo, i gerghi – se devono essere criptici (ma si v. in merito almeno Vigolo 2010; Marcato 2013: 10; Sanga 2015) – devono esserlo principalmente per quei parlanti con cui sono maggiormente a contatto¹³.

3.2. La struttura dei nomi e degli aggettivi

A riguardo della struttura di nomi e aggettivi, i meccanismi di formazione di maschili e femminili e di singolari e plurali appaiono identici rispetto alle varietà ospiti: non si apprezza, perciò, alcun aumento o

¹² Per completezza, anche i gerghi che parassitano una varietà a sola negazione preverbale non cambiano la struttura della frase negativa ospite: si v. il tarón di Tuenno e Val di Non: *slàkiɽe ke se no'l desmet el zoi en čamoča* ‘digli che se non smette lo mando in prigione’ (Menapace 2019: 116).

¹³ Un’analisi di grana più fine, che non mi è permesso portare a compimento qui, potrà mostrare eventuali specializzazioni e usi pragmatici dei negatori *miga* e *vita* in rügín o di *brisa* e *nekka* nei gerghi bolognesi. Si tratta comunque di percorsi di grammaticalizzazione per nulla eccezionali, comparabili, per restare al dominio gallo-italico, a quello abbozzato da Ballarè/Goria (2019) per i negatori *nen* e *pa* che appaiono in competizione nel torinese del Sei e Settecento.

diminuzione di classi morfologiche in questo microsistema. Forse gli unici significativi cambiamenti sul piano morfologico hanno a che fare con la propensione alla derivazione di nomi, o di aggettivi che poi vengono frequentissimamente sostantivati, tramite la suffissazione di morfemi dall'aspetto germanico oppure discendenti dal latino (Sanga 1989: 20-21). Anche se tramite l'aggiunta di suffissi il numero di morfemi presenti nelle parole gergali aumenta rispetto alle parole delle varietà ospite, ciò non dà comunque mai luogo a un aumento di classi o subclasse flessive.

Se, poi, prendiamo in considerazione i gerghi meccanici, e in particolare quelli a inversione, noteremo che le sillabe che compongono ciascuna parola non vengono mai aumentate, ma piuttosto scambiate di posizione attraverso procedimenti più o meno semplici. Questi procedimenti meccanici evitano di fatto che si realizzi la flessione per la maggior parte dei nomi e degli aggettivi, che diventano quindi invariabili per numero, mentre le varietà ospiti, almeno per la maggior parte dei femminili, continuano a distinguere singolare e plurale. Si vedano gli esempi successivi, presi dalla damo da ntradi (10a), dal larpa iudre (gergo usato non solo 'per gioco', ma anche come normale veicolo di comunicazione nell'ambiente postale del mendrisiotto, (11a))¹⁴ e dai dialetti che li ospitano (10b-11b).

(10) a. Damo da ntradi

al lamsa – aj lamsa, al pelca – aj pelca, la tabu – aj tabu

'il salame – i salami, il cappello – i cappelli, la bottiglia – le bottiglie' (Cena 2006: 34)

(10) b. Ciglianese

al salam – aj salam, al capel – aj capej, [la='buta] – [le='bute]

'il salame – i salami, il cappello – i cappelli, la bottiglia – le bottiglie' (*ibid.*, con modifiche)

(11) a. Larpa iudre

néca – i néca, dèlfra – i dèlfra, tanapü – i tanapü

'cane – i cani, fratello – i fratelli, puttana – le puttane' (Lurà 1987: 218-227)

(11) b. Mendrisiotto

can – can, fradèll / fredèll – fredèi, pütana – pütani

¹⁴ Per la damo, v. Berruto/Vicari (2008-2009: 20-23). Il larpa iudre è stato descritto da Berruto (1979-1980) e da Lurà (1987: 221-226).

‘cane – i cani, fratello – fratelli, puttana – puttane’ (Lurà 1987: *passim*)

Pur se non mancano le eccezioni in entrambi i sistemi gergali, sia nella damo da ntradi sia nel larpa iudre «[p]er la formazione del plurale dei sostantivi, generalmente viene mantenuta invariabile la forma singolare, con l’assegnazione di numero affidata all’articolo» (Berruto/Vicari 2008-2009: 23). Si ha pertanto una riduzione delle classi flessive e delle categorie morfologiche marcate sui nomi e sugli aggettivi, ovvero, secondo le proposte di Trudgill, un caso semplificazione dei paradigmi del gergo¹⁵.

4. Conclusioni

Tentiamo ora di raccogliere i dati discussi sin qui in un quadro generale, dal quale poi trarre qualche conclusione. Questa indagine ha forse il limite di essere stata svolta a campione su alcuni microsistemi grammaticali, ma è pur vero che semplificazione e complessificazione si valutano, di norma, quantomeno in sociolinguistica, sempre su un dato microsistema o una data struttura (Berruto 2012: 91; v. anche il metodo di Trudgill, e.g. 2009; 2011). Non sembra però che le semplificazioni riscontrate portino con sé ristrutturazioni complicanti in altri domini o microsistemi: si veda ad esempio l’illustrazione del sistema pronominale tonico (§ 3.1), la cui ristrutturazione dà adito a una possibile ulteriore semplificazione (e non complessificazione) nella coniugazione verbale.

Comunque sia, l’ipotesi iniziale, chiamata sopra la “congettura di Trudgill”, voleva che i gerghi risultassero almeno un po’ più complessi delle loro varietà ospiti, a causa del fatto che le comunità di gerganti sono piccole, molto coese e con un ampio ventaglio di conoscenze condivise. Quest’ipotesi – si sarà capito – non trova piena conferma in quanto abbiamo visto nei paragrafi precedenti.

Il paradigma dei pronomi personali di molti gerghi gallo-italici è meno (o comunque ugualmente) numeroso rispetto al dialetto che li ospita, anche se il numero di morfemi per ciascuna singola unità tende

¹⁵ Si aggiunga che nel microsistema gergale si ha anche l’eliminazione di ridondanza in casi come *aj capej* di contro a *aj pelca*, che dall’occhio del tipologo potrebbero essere analizzati come segue: *a-j cape-j* ‘DEF-PL cappello-PL’, con doppia marca di plurale; *a-j pelca* ‘DEF-PL cappello’, con una sola marca di plurale.

ad essere maggiore. Nel dominio della negazione di frase generalmente le strutture di gergo e varietà ospite rimangono tipologicamente identiche, senza che vi sia l'eliminazione di strutture ridondanti o, al contrario, complessificazione di strutture semplici. Infine, se anche si può dire che i nomi e gli aggettivi gergali possano avere un maggior numero di morfemi rispetto ai dialetti ospitanti e maggiore opacità lessicale per i parlanti di questi ultimi, si è messo in luce che non è raro, almeno nei gerghi cosiddetti a inversione, avere una diminuzione delle categorie morfologiche dei paradigmi. In ogni caso, non si verificano mai casi in cui nella varietà gergale si riscontra la creazione di nuove categorie morfologiche (e conseguente eliminazione di marche ridondanti, v. nota 14). Nell'insieme, almeno per i domini che abbiamo esaminato, i gerghi sono più semplici delle loro varietà ospite.

Tra i parametri rilevanti nella "congettura di Trudgill" abbiamo trascurato il contatto con altre varietà linguistiche, in quanto i gerghi, a differenza delle *societies of intimates* tipiche, vengono praticati da comunità di parlanti che sono perlopiù itineranti e girovaghe, e quindi entrano frequentemente in contatto con altre lingue e varietà. Il ruolo che gioca il contatto nella complessificazione o semplificazione dei sistemi linguistici può essere ambivalente, ma le comunità gerganti non sono certo composte principalmente o esclusivamente di persone che vi entrano e imparano il gergo da adulte: specialmente in anni recenti anzi sarà vero piuttosto che il gergo viene appreso in famiglia o dagli adulti che si frequentano fin dalla tenera età (si pensi alle famiglie di giostrai). Questa seconda condizione dovrebbe portare a processi di complessificazione (Trudgill 2010: 314). A maggior ragione, dunque, occorre riformulare almeno parzialmente la "congettura di Trudgill" sulla scorta dei risultati della nostra analisi. Ad essi si confà bene, a quanto mi pare, la proposta avanzata, in particolare nel merito della tipologia fonetica, da Schreier (2016: 145): quando il contatto coinvolge lingue tipologicamente diverse, allora esso porta necessariamente alla semplificazione. Se invece il contatto avviene tra varietà di una stessa lingua, con sistemi simili, allora l'esito può essere sia la complessificazione sia la semplificazione. È evidente che i gerghi gallo-italici sono prevalentemente parlati in situazioni nelle quali si verifica il secondo tipo di contatto; si prenda ad esempio il rügin: il gergo è a tutti gli effetti una varietà (diafasico-diastratica) del dialetto

gallo-italico che lo ospita; così, *variatis variandis*, la damo da ntradi e il larpa iudre ecc.

La passeggiata tipologico-sociolinguistica compiuta tra i gerghi dell'Italia settentrionale e della Svizzera italiana ci ha permesso se non altro di toccare con mano quanto queste varietà, forse ormai quasi estinte, certamente un po' trascurate al di là degli studi specifici e della dialettologia, possano essere utili per la ricerca linguistica, anche teorica. Di certo, per proseguire su questa strada, oltre che saggiare su altri domini grammaticali le parziali conclusioni cui siamo pervenuti in questo articolo, non sarà inutile procedere a nuove raccolte di dati, svolte il più possibile in situazioni autentiche, ora che alcuni degli ultimi gerghi sono ancora in vita.

Riferimenti bibliografici

- AIS = Jud, Jacob & Jaberg, Karl. 1928-40. *Sprach- und Sachatlas Italiens und der Sudschweiz*, 8 voll. Zofingen: Ringer u.C.
- Aly Belfadel, Arturo. 1898. Sopra un gergo di commessi di negozio torinesi. *Archivio di psichiatria, scienze penali ed antropologia criminale* 11. 633-636.
- Arigoni, Nicola. 2020. L'altro altro. Prestiti nel gergo dei calderai ambulanti della Val Colla (Canton Ticino). *Argotica* 1(9). 21-34.
- Arigoni, Nicola & Vicari, Mario (a cura di). 2019. *Documenti orali della Svizzera italiana. Capriasca, seconda parte*. Bellinzona: Centro di dialettologia e di etnografia.
- Ballarè, Silvia & Gorla, Eugenio. 2019. Qual è il contributo dei dialetti per lo studio diacronico della negazione? Il caso del torinese. In Gianollo, Chiara & Mauri, Caterina (a cura di), *CLUB Working Papers in Linguistics* 3, 208-216. Bologna: Circolo Linguistico dell'Università di Bologna.
- Berruto, Gaetano. 1979-1980. Una lingua per l'occasione a Mendrisio: il "larpa iudre". *Bollettino dell'Atlante Linguistico Italiano*, n.s., III(3-4). 87-89.
- Berruto, Gaetano. 2012. *Sociolinguistica dell'italiano contemporaneo*. Roma: Carocci (prima ed. Roma, La Nuova Italia Scientifica, 1987).
- Berruto, Gaetano. 1990. Semplificazione linguistica e varietà sub-standard. In Holtus, Günter & Radtke, Edgar (Hrsg.), *Sprachlicher Substandard*

- III. *Standard, Substandard und Varietätenlinguistik*, 17-43. Tübingen, Niemeyer.
- Berruto, Gaetano & Vicari, Mario. 2008-2009. Di due gerghi a modificazione sillabica in area galloitalica. *LIDI* II(3-4). 7-32.
- Bertolotti, Guido, Bralla, Felice, Butti, Carlo & Sanga, Glauco. 1978. I magnani della Val Cavargna e il loro gergo. In Leydi, Roberto & Sanga, Glauco (a cura di), *Como e il suo territorio*. 373-464. Milano: Silvana.
- Bossong, Georg. 2016. Classifications. In Ledgeway, Adam & Maiden, Martin (eds), *The Oxford Guide to the Romance Languages*, 63-72. Oxford & New York: Oxford University Press.
- Cena, Umberto. 2006. *La damo da ntradi: un gergo a Cigliano (VC)*. Torino: Università degli Studi. (Tesi di laurea triennale.)
- Cherubini, Francesco. 1843. *Vocabolario milanese-italiano*. IV vol. Milano: Imperiale Regia Stamperia.
- Collins, Chris & Moody, Simanique & Postal, Paul M. 2008. An AAE Camouflage Construction. *Language* 84(1). 29-68.
- Cortelazzo, Manlio. 1979. Una parola veneta antica e rara: 'meòden(a)'. In Aa.Vv., *Medioevo e rinascimento veneto con altri studi in onore di L. Lazzarini*, vol. II: 323-325. Padova: Antenore.
- Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity*. Amsterdam & Philadelphia: John Benjamins.
- Daniel, Michael. 2013. Plurality in Independent Personal Pronouns. In Dryer, Matthew S. & Haspelmath, Martin (eds), *The World Atlas of Language Structures Online*, <http://wals.info/chapter/35>. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Diamond, Jared. 2012. *The World until Yesterday*. London: Penguin.
- Dryer, Matthew S. 2013. Negative Morphemes. In Dryer, Matthew S. & Haspelmath, Martin (eds), *The World Atlas of Language Structures Online*, <http://wals.info/chapter/112>. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Ferrero, Ernesto. 1991. *Dizionario storico dei gerghi italiani: dal Quattrocento a oggi*. Milano: Mondadori.
- Fiorentino, Giuliana. 2005. Dialetti in rete. *Rivista italiana di dialettologia* 29. 111-149.
- Fiorentino, Giuliana. 2009. Complessità linguistica e variazione sintattica. *Studi italiani di linguistica teorica e applicata* XXXVIII(2). 281-312
- Fiorentino, Giuliana. 2019. Complessità e semplicità come qualità di un sistema linguistico: codici artificiali e codici naturali a confronto. In

- Burov, Ivaylo & Fiorentino, Giuliana (dir.), *Complexité des structures et des systèmes linguistiques: le cas des langues romanes*, 37-57. Sofia: CU Romanistika.
- Harris, Alice. 1981. *Georgian Syntax*. Cambridge: Cambridge University Press.
- Keller, Oscar. 1933. Due testi in dialetto di Maglio di Colla. *L'Italia dialettale* 9. 214-229.
- Keller, Oscar. 1934. Die Geheimsprache der wandernden Kesselflicker der Val Colla, Tessin. *Völkstum und Kultur der Romanen* 7(1). 55-81.
- Keller, Oscar. 1937. Die Mundarten des Sottoceneri (Tessin) dargestellt an Hand von Paralleltexen. II, Lugano und das Basso Luganese. *Revue de linguistique romane* 13. 127-361.
- Kusters, Wouter. 2003. *Linguistic complexity. The influence of social change on verbal inflection*. Utrecht: LOT.
- Kornai, András. 2013. Digital Language Death. *PLoS ONE* 8(10): e77056. <https://doi.org/10.1371/journal.pone.0077056>
- Iannàccaro, Gabriele. 2000. Per una semantica più puntuale del concetto di dato linguistico: un tentativo di sistematizzazione epistemologica. *Quaderni di Semantica* 2(1). 51-79.
- Lurà, Franco. 1987. *Il dialetto del Mendrisiotto. Descrizione sincronica e diacronica e confronto con l'italiano*. Mendrisio & Chiasso: Edizione Unione di Banche Svizzere.
- Lurati, Ottavio. 1990. Quale l'ideologia degli ambulanti? Il gergo dei magnani lombardi con una raccolta inedita sulla Val Colla della metà dell'Ottocento. In Zappa, Fernando (a cura di), *Valli di Lugano*, 221-248. Locarno: Dadò.
- Manni, Eugenio. 1973. *Il tariùsc. La parlata degli ombrellai*, Varallo Sesia: Arti Grafiche Valsesiane.
- Marcato, Carla. 2013. *I gerghi italiani*. Bologna: il Mulino.
- McWhorter, John. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5. 125-166.
- Menapace, Remo. 2019. *Il Tarón degli spazzacamini*. Abano Terme: Aldo Francisci.
- Menarini, Alberto. 1942. *I gerghi bolognesi*. Modena: Società tipografica modenese.
- Miestamo, Matti & Sinnemäki, Kaius & Karlsson, Fred (a cura di). 2008. *Language complexity: Typology, contact, change*. Amsterdam: John Benjamins.

- Moretti, Bruno. 2017. Che cosa ha da dire la sociolinguistica sul tema della complessità delle lingue. *Rivista Italiana di Dialettologia* 41. 35-52.
- Mufwene, Salikoko S. & Coupé, Christophe & Pellegrino, François (a cura di). 2017. *Complexity in language. Developmental and evolutionary perspectives*. Cambridge: Cambridge University Press.
- Papanti, Giovanni. 1875. *I parlari italiani in Certaldo*. Livorno: Francesco Vigo.
- Patrucco, Elisa. 2003. Sul dialetto in Internet. *Rivista Italiana di dialettologia* 27. 139-174.
- Prati, Angelico. 1978. *Voci di gerganti, vagabondi e malviventi studiate nell'origine e nella storia*. Nuova edizione con una nota bibliografica e una postilla critica di Tristano Bolelli. Pisa: Giardini (prima ed. Pisa, 1940).
- Rohlf, Gerhard. 1969. *Grammatica storica della lingua italiana e dei suoi dialetti. Sintassi e formazione delle parole*. Torino: Einaudi.
- Salvioni, Carlo. 1891. La gita di un glottologo in Val Colla (agosto 1890). *Bollettino storico della Svizzera italiana* XIII. 94-196.
- Sampson, Geoffrey & Gil, David & Trudgill, Peter (a cura di). 2009. *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press
- Sanga, Glauco. 1977. Il gergo dei magnani della val Cavargna e i gerghi italiani. In Butti, Carlo (a cura di), *Il 'rugin'*, 7-9. Cavargna: Associazione 'Amici di Cavargna'.
- Sanga, Glauco. 1984. *Dialettologia lombarda*. Pavia: Università di Pavia.
- Sanga, Glauco. 1989. Estetica del gergo. Come una cultura si fa forma linguistica. *La Ricerca Folklorica* 19. 17-26.
- Sanga, Glauco. 1993. Gerghi. In Sobrero, Alberto A. (a cura di), *Introduzione all'italiano contemporaneo. La variazione e gli usi*, vol. I: 151-189. Bari: Laterza.
- Sanga, Glauco. 2015. La segretezza del gergo. In Cugno, Federica & Rivoira, Matteo & Specchia, Maria Sabrina (a cura di), *Studi linguistici in onore di Lorenzo Massobrio*, 884-903. Alessandria: dell'Orso.
- Scala, Andrea. 2016. Gerghi storici nell'Italia settentrionale odierna. In Raus, Rachele & Bălă, Laurentiu (a cura di), *Sul gergo nel XXI secolo. Despre argou în secolul XXI. Sur l'argot au XXIe siècle. Collecția Argotolog* 1, 47-58. Craiova: Editura Universitaria Craiova.
- Schreier, Daniel. 2016. A true *split*? Typological and sociolinguistic considerations on contact intensity effects. In Baechler, Raffaella & Seiler, Guido (eds), *Complexity, Isolation, and Variation*, 139-157. Berlin & Boston: Mouton de Gruyter.

- Siewierska, Anna. 2013. Gender Distinctions in Independent Personal Pronouns. In Dryer, Matthew S. & Haspelmath, Martin (eds), *The World Atlas of Language Structures Online*, <http://wals.info/chapter/44>. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Soldati, Felice. 1957. Glossarietto del “rugin”. Storia del figliuol prodigo. *Svizzera italiana* 17(125). 6-14.
- Spears, Arthur. 1998. African-American language use: Ideology and so-called obscenity. In Mufwene, Salikoko S. & Rickford, John R. & Bailey, Guy & Baugh, John (eds), *African American English: Structure, history, and usage*, 226-250. New York: Routledge.
- Spitzer, Leo. 2014. *Lettere di prigionieri di guerra italiani 1915-1918*. 2a ed. Torino: Bollati Boringhieri.
- Tomasini, Renzo. 1989. *Il dialetto della Val Rendena*. S. Michele all'Adige: Museo della Gente Trentina.
- Toso, Fiorenzo. 2002. Liguria. In Cortelazzo, Manlio & Marcatò, Carla & De Blasi, Nicola & Clivio, Gianrenzo P. (a cura di), *I dialetti italiani. Storia, struttura, uso*, 196-225. Torino: UTET.
- Traina, Antonino. 1868. *Nuovo vocabolario siciliano-italiano*. Palermo: Giuseppe Pedone Lauriel.
- Trudgill, Peter. 2009. Sociolinguistic typology and complexification. In Sampson *et al.* (2009), 98-109.
- Trudgill, Peter. 2010. Contact and Sociolinguistic Typology. In Hickey, Raimond (ed.), *The Handbook of Language Contact*, 299-319. Malden, MA: Wiley-Blackwell.
- Trudgill, Peter. 2011. *Sociolinguistic typology*. Oxford: Oxford University Press.
- Trudgill, Peter. 2012. On the sociolinguistic typology of complexity loss. *Language Documentation & Conservation*, special issue 3. 90-95.
- Trudgill, Peter. 2015. Societies of intimates and linguistic complexity. In De Busser, Rik & La Polla, Randi J. (eds), *Language Structure and Environment*, 133-148. Amsterdam & Philadelphia: John Benjamins.
- Vigolo, Maria Teresa. 2010. Gergo. In Simone, Raffaele (dir.), *Enciclopedia dell'italiano*, 565-567. Roma: Istituto della Enciclopedia Italiana.
- Vigolo, Maria Teresa & Barbierato, Paola. 2008. Il gergo storico e l'uso del nome proprio. In D'Achille, Paolo & Caffarelli, Enzo (a cura di), *Lessicografia e Onomastica 2, Atti delle giornate internazionali di Studio, Università degli Studi Roma Tre - 14-16 febbraio 2008*, 361-372. Roma: Società editrice romana.

- Vitali, Daniele. 2020. *Dialetti emiliani e dialetti toscani*. 4 voll. Bologna: Pendragon.
- Wagner, Max Leopold. 1990. *Sondersprachen der Romania II. Sardinien, Italien, Portugal, Rumänien und Türkei*. Stuttgart: Franz Steiner.
- Wright, Sue. 2006. Regional or minority languages on the WWW. *Journal of Language and Politics* 5(2). 189-216.

ADRIANO MURELLI

Strategie di relativizzazione nelle lingue europee: tra tipologia e sociolinguistica

This paper explores possible interactions between typology and sociolinguistics by taking relativization strategies in standard and non-standard varieties in a sample of thirty-six European languages as an example. A purely typological-functional analysis of the opposition between strategies attested in the two groups of varieties does not provide satisfactory results. A perspective shift towards sociolinguistics allows us to identify principles and criteria underlying the selection of strategies to be considered standard during the codification process and to formulate an implicational scale of standardness representing the reciprocal relations between the strategies. The validity of the scale of standardness is critically discussed, pondering its strengths and weaknesses.

Parole chiave: frasi relative, lingue europee, non-standard, tipologia, sociolinguistica.

1. Introduzione

Questa ricerca è nata sulla scorta di *Dialectology meets typology* (Kortmann 2004), volume miscelaneo che si proponeva di indagare la variazione sociolinguistica – in particolare quella diatopica – in prospettiva tipologico-funzionale. Analizzando le lingue europee, caratterizzate solitamente da una precisa tradizione grammaticale e dall'esistenza di una varietà standard codificata, i lavori tipologici si sono perlopiù concentrati su quest'ultima e hanno dedicato un'attenzione molto minore a varietà lontane dal centro dell'architettura della lingua. È merito dunque di studi come Kortmann (2004) aver tematizzato la necessità di indagare, oltre alla varietà di riferimento di una lingua, anche varietà marcate in diatopia, diastratia e diafasia¹.

¹ Kortmann (2004) non è il primo a sottolineare l'importanza di occuparsi della grammatica delle varietà non-standard. Già Glaser (1997) – riferendosi specialmente

Idealmente situato all'incrocio degli assi sociolinguistici (Berruto 2012: 24), lo standard costituisce una varietà "neutra", risultato di operazioni di livellamento e compromesso (cfr. par. 3). Se i dati di uno studio vengono tratti esclusivamente da varietà standard, il quadro tipologico che ne emerge rischia di non essere equilibrato: la variazione interna viene obliterata e i tratti della varietà standard sono automaticamente identificati come tratti caratteristici della lingua in oggetto. Le strutture presenti nello standard vengono quindi messe in risalto a scapito di quelle documentate solo in varietà non-standard. Un simile squilibrio e le eventuali distorsioni che ne derivano rispetto a un quadro tipologico globale sono difficilmente sanabili, a meno di non dichiarare esplicitamente che i risultati ottenuti valgono esclusivamente per le varietà standard delle lingue considerate.

Scopo di questo contributo è mostrare come sia possibile coniugare tipologia e sociolinguistica servendosi degli strumenti di entrambe le discipline per delineare un quadro complessivo della variazione linguistica nell'area europea. Il fenomeno scelto per l'analisi sono le strategie di relativizzazione attestate in un campione di trentasei lingue europee. Si adotterà una doppia prospettiva: la comparazione interlinguistica sarà effettuata sulla base di dati tratti non solo dalla varietà standard, ma anche da altre varietà che compongono il diasistema delle lingue del campione. Per rendere gestibile la complessità di un simile proponimento, è parso utile adottare una definizione operativa del termine 'non-standard': partendo dalla considerazione che la varietà standard è considerata neutra, ossia non marcata, situata al centro dell'architettura della lingua, l'etichetta di 'non-standard' è stata attribuita a tutte le varietà marcate sociolinguisticamente su uno degli assi di variazione.

Il contributo è organizzato come segue: nel par. 2 si illustrerà come un approccio puramente tipologico-funzionale non sia in grado di rendere conto compiutamente della distribuzione delle strategie

al tedesco – invita a esaminare più da vicino la sintassi dialettale, ambito di ricerca che un progetto di ampio respiro come EURO TYP (1990-1994) – per quanto incentrato sull'area linguistica europea, le cui lingue componenti presentano diasistemi articolati – aveva purtroppo trascurato (van der Auwera 2011). I contributi in Kortmann (2004) hanno il merito di analizzare fenomeni riscontrati nelle varietà non-standard di singole lingue europee da una prospettiva tipologico-funzionale, oltre ad affrontare questioni teoriche quali l'esistenza di universali vernacolari.

di relativizzazione tra le varietà standard e non-standard delle lingue europee considerate; nel par. 3 si adotterà dunque una prospettiva sociolinguistica, identificando una serie di criteri che possono essere stati alla base della classificazione di diverse strategie come standard o non-standard. Quindi, nel par. 4 si introdurrà una scala di standardità per (alcune del)le strategie di relativizzazione emerse dall'analisi tipologica, osservando come le lingue del campione si pongono rispetto alla scala stessa. Infine, nel par. 5 si sottoporrà la scala di standardità a un esame critico, mettendone in luce punti di forza e di debolezza, e si discuteranno questioni concernenti la sua validità, anche nel quadro della formazione di varietà neo-standard.

2. Dalla prospettiva tipologica...

Uno studio tipologico delle strategie di relativizzazione nell'area linguistica europea, con particolare attenzione alle varietà non-standard, è condotto in Murelli (2011). Murelli esamina i dati di un campione di trentasei lingue europee² traendoli da grammatiche di riferimento, contributi scientifici in volume o in rivista e, in misura minore, da questionari e corpora. Per classificare le strategie di relativizzazione – ossia le strategie utilizzate per la codifica morfosintattica di frasi relative³ – Murelli adotta tre parametri comunemente impiegati nella

² Le lingue sono le seguenti (tra parentesi l'abbreviazione usata nelle tabelle, che segue, con una eccezione, lo standard ISO 639-2): albanese (ALB), basco (BAQ), bielorusso (BEL), bosniaco/croato/serbo (BCS), bulgaro (BUL), catalano (CAT), ceco (CZE), danese (DAN), estone (EST), finlandese (FIN), francese (FRE), greco (GRE), inglese (ENG), irlandese (GLE), islandese (ICE), italiano (ITA), lettone (LAV), lituano (LIT), macedone (MAC), maltese (MLT), norvegese (NOR), olandese (DUT), polacco (POL), portoghese (POR), rumeno (RUM), russo (RUS), slovacco (SLK), sloveno (SLV), sorabo inferiore (DSB), sorabo superiore (HSB), spagnolo (SPA), svedese (SWE), tedesco (GER), turco (TUR), ucraino (UKR), ungherese (HUN).

³ Per gli scopi di questo saggio, sulla scorta di Lehmann (1984) e Zifonun (2001) possiamo definire una frase relativa come una frase subordinata con funzione di modificatore di una struttura nominale sovraordinata (detta antecedente o testa) contenuta in una frase matrice. La frase relativa può o individuare il referente della struttura nominale in una schiera di possibili referenti o fornire informazioni aggiuntive su di esso. Nel primo caso la frase relativa è detta restrittiva, nell'altro non-restrittiva o attributiva.

letteratura sul tema (Lehmann 1984, de Vries 2002, Andrews 2007, Cristofaro & Giacalone Ramat 2007):

1. l'ordine frase matrice-frase relativa, differenziando strategie incasate e giustapposte e classificando le prime in prenominali, postnominali e circumnominali, le seconde in anteposte e posposte;
2. l'elemento relativo, che codifica il componente relativizzato nella frase relativa, individuando elementi semplici (composti da una sola unità morfosintattica) ed elementi combinati (composti da più unità morfosintattiche) e distinguendo, tra i primi, pronomi relativi, elemento relativo specializzato, particella relativa e marca zero (Tabella 1);
3. le posizioni sintattiche che una strategia può relativizzare, rifacendosi alla Gerarchia di Accessibilità (*Accessibility Hierarchy*, Keenan & Comrie 1977).

Tabella 1 - *Gli elementi relativi semplici e le proprietà che codificano*⁴

	1	2	3	4	5	6
Elemento relativo:	RPRO	RPRO	RPRO	SRE	RPAR	Ø
Codifica di:						
legame FM-FR	+	+	+	+	+	-
ruolo sintattico	+	+	+	+(∃!)	-	-
accordo di genere	+	+	-	-	-	-
accordo di numero	+	-	-	-	-	-

Partendo dalle proposte di Lehmann (1984) e Cristofaro & Giacalone Ramat (2007), Murelli (2011: 85) individua quattro proprietà che un elemento relativo può codificare nelle lingue del campione: legame tra frase matrice e frase relativa, ruolo sintattico del componente relativizzato nella frase relativa, accordo di genere e accordo di numero con la testa nominale. Come illustrato nella Tabella 1, se un elemento codifica almeno le prime due proprietà è classificato come pronome relativo (colonne 1-3)⁵; se codifica solo il legame tra frase matrice e

⁴ Le abbreviazioni utilizzate nelle tabelle e negli esempi sono riportate in calce al saggio, prima dei riferimenti bibliografici.

⁵ A seconda delle proprietà che codificano, è possibile suddividere i pronomi relativi in flessi per genere e numero, flessi per genere e non flessi per genere e numero. Tra i primi abbiamo, per esempio, il pronome relativo *který* del ceco, declinato per tre generi e due numeri; tra i secondi l'inglese *who/which*, in cui la forma selezionata

frase relativa è definito particella relativa (colonna 5). La marca zero (colonna 6), presente ad esempio in inglese (*The woman [Ø we met yesterday] was an old friend of mine*), non codifica nessuna proprietà, mentre l'elemento relativo specializzato (colonna 4) codifica le prime due proprietà, come il pronome relativo, ma può relativizzare solo una posizione sintattica (fenomeno indicato con il simbolo 'Ξ!' nella Tabella 1), come l'elemento *cuyo/cuya* in spagnolo, impiegabile solo per la posizione di genitivo. In (1)-(3) troviamo istanze di pronome relativo, elemento relativo specializzato e particella relativa. L'elemento relativo è in grassetto, la frase relativa è posta tra parentesi quadre.

- (1) ceco
Člověk, [který nekouří,]
 uomo RPRO.NOM.SG.M NEG.fuma
ušetří.
 risparmia
 'Una persona che non fuma risparmia.' (Petr 1987: 524)
- (2) polacco
Przewodnik pokazal nam pomieszczenia,
 guida ha.mostrato a.noi sale
[gdzie odbywa się remont.]
 SRE procede REFL ristrutturazione
 'La guida ci ha mostrato le sale dove stanno effettuando lavori di ristrutturazione.' (da questionario)
- (3) portoghese
Ele era um homem [que você]
 lui era DET uomo RPAR tu
se dava bem.]
 ti.rapportavi bene
 'Era un uomo con cui si andava d'accordo.' (da questionario)

è legata alla distinzione tra antecedenti [\pm umano]/[\pm animato], qui assimilata, sulla scorta di Cristofaro & Giacalone Ramat (2007), a quella di genere; tra gli ultimi il francese *qui*, forma che non segnala accordo di genere e numero con il referente. A sua volta, il ruolo sintattico del componente relativizzato può essere veicolato ora tramite morfemi flessivi (polacco *książka, któr-ą czytam* 'il libro che (RPRO.ACC.SG.F) sto leggendo', ruolo di oggetto diretto), ora tramite la combinazione con adposizioni (italiano *l'amico con il quale stavo parlando*, ruolo di obliquo), ora tramite una combinazione di entrambi (polacco *przyjaciel, z któr-ym rozmawiałem* 'l'amico con il quale (RPRO.INSTR.SG.M) stavo parlando', ruolo di obliquo).

In (4)-(6), invece, sono riportati esempi di elementi relativi combinati – rispettivamente, pronomi relativi con particella relativa, pronomi relativi con elemento (clitico) di ripresa e particella relativa con elemento di ripresa.

- (4) tedesco (dialetto alemannico)

D Kirch, [neue dere wu
 DET chiesa accanto RPRO.DAT.SG.F RPAR
er wohnt,] isch im Griech kabütt gemacht wor.
 lui vive è in.DET guerra a.pezzi fatto stato
 ‘La chiesa accanto alla quale vive è stata distrutta durante la guerra.’ (Balliet 1997: 214)

- (5) catalano

uns marrecs [a qui la Maria
 alcuni ragazzi a RPRO DET Maria
els dóna galetes]
 CL.DAT.3PL dà biscotti
 ‘Alcuni ragazzi a cui Maria dà biscotti.’ (Hualde 1992: 58)

- (6) bulgaro

Sega imam chimikalka, [deto moga da
 ora ho penna RPAR posso COMP
piša s neja s časove.]
 scrivo con lei da ore
 ‘Ora ho una penna con cui posso scrivere per ore.’ (da questionario)

L’indagine delle varietà non-standard delle trentasei lingue europee incluse nel campione, condotta in Murelli (2011: cap. 4), mette in luce la presenza di diverse strategie di relativizzazione non considerate o considerate solo marginalmente in studi tipologici e di tipologia areale dedicati a questo tema (ad esempio Zifonun 2001, Cristofaro & Giacalone Ramat 2007, Comrie & Kuteva 2013). In particolare, il parametro ‘elemento relativo’ mostra una notevole variabilità: accanto agli elementi relativi semplici compaiono diversi elementi combinati. La Tabella 2, ripresa da Murelli (2011: 244), elenca in ordine di frequenza gli elementi relativi individuati nelle lingue del campione e le proprietà che codificano. La colonna ‘n’ riporta il numero di lingue in cui ciascun elemento ricorre.

Mentre alcune delle strategie sono attestate solo in varietà non-standard – per esempio buona parte di quelle basate sugli elementi relativi riportati nelle posizioni più basse della Tabella 2 –, ve ne sono altre, tipologicamente identiche tra loro, che compaiono nelle varietà standard di alcune lingue e nelle varietà non-standard di altre. Un caso tipico è la combinazione tra una particella relativa e un elemento di ripresa: presente in trentuno lingue su trentasei, è considerata ora standard, ora non-standard; non solo, anche in prospettiva intralinguistica diverse grammatiche della stessa lingua contengono a volte giudizi contrastanti circa il suo grado di standardità (cfr. più nel dettaglio il par. 4).

Muovendosi all'interno dell'approccio tipologico tradizionale, Murelli (2011: 251ss.) prova a spiegare la presenza di determinate strategie nelle varietà standard o non-standard delle lingue del campione ricorrendo a principi funzionali (iconicità, analogia, economia). Tuttavia, questa operazione non porta i risultati sperati: non è possibile argomentare in maniera convincente che le strategie attestate nelle varietà non-standard seguano principi funzionali diversi rispetto a quelle attestate nelle varietà standard. Un esempio: la strategia del pronome relativo – diffusa specialmente, ma non solo, nelle varietà standard – è più economica rispetto a quella della strategia della particella relativa con pronome di ripresa, attestata tanto in varietà standard quanto in varietà non-standard, poiché codifica diverse proprietà morfosintattiche tramite una sola forma; per lo stesso motivo è però meno iconica dell'altra (in alcuni casi non codifica significati distinti tramite morfemi distinti). Viceversa, l'uso (peculiare, seppure non limitato al non-standard) di particelle relative che hanno spesso carattere polivalente – oltre a relazioni relative possono codificare altri tipi di relazioni subordinate, come il *che* polivalente italiano – può essere visto come non-iconico, ma è economico. Pare dunque possibile ricorrere agli stessi principi per giustificare la presenza di strategie in entrambi i gruppi di varietà, senza ottenere chiare demarcazioni.

Tabella 2 - *Gli elementi relativi attestati nelle lingue del campione e le proprietà che codificano*

#	Proprietà codificate Elemento relativo	Legame FM-FR	Ruolo sintattico	Accordo genere	Accordo numero	n
1	RPAR	+	-	-	-	34
2	RPRO	+	+	+/-	+/-	32
3	SRE	+	+	-	-	31
	RPAR+RE	+	+	+/-	+/-	31
4	RPRO+RE	+	+	+/-	+/-	11
5	SRE+RE	+	+	+/-	+/-	8
6	Ø	-	-	-	-	6
7	Ø+RE	-	+	+/-	+/-	5
8	RPRO+RPAR	+	+	+/-	+/-	5
	SRE+RPAR	+	+	-	-	5
9	RPAR+PAR	+	-	-	-	3
10	RPAR+RPRO	+	+	+/-	+/-	2
	RPAR+SRE	+	+	-	-	2
11	RPRO+PAR	+	+	+/-	+/-	1
	Ø+PAR	-	-	-	-	1
	RPRO+RPAR+RE	+	+	+/-	+/-	1
	RPAR+PAR+RE	+	+	+/-	+/-	1
	SRE+RPAR+RE	+	+	+/-	+/-	1

D'altro canto, esistono strategie che contrastano con i principi funzionali: diversi elementi combinati (tra cui quelli che occupano le posizioni 4-5 e 8-11 nella Tabella 2, attestati perlopiù, ma non esclusivamente, in varietà non-standard) sono antieconomici, poiché codificano più volte gli stessi parametri. Ad esempio, un pronome relativo combinato a una particella relativa codifica due volte il legame tra frase matrice e frase relativa; un pronome relativo combinato a un pronome di ripresa (attestato anche in varietà standard) codifica due volte il ruolo sintattico del componente relativizzato ed eventualmente l'accordo di genere e numero. Viceversa, l'uso di marche zero può essere considerato anti-iconico, perché viene meno la corrispondenza forma-funzione, mancando la prima del tutto.

Di fronte a questa *impasse*, si rivela utile riconsiderare la distinzione 'standard vs. non-standard': essa non si applica alle strategie di relativizzazione, che da questo punto di vista sono neutre, in quanto caratterizzate da una precisa struttura descrivibile tramite parametri tipologici, ma alle varietà in cui le strategie sono attestate. A sua volta, la classificazione di una varietà come standard o non-standard non è intrinseca alla varietà stessa, ma avviene ad opera di istanze esterne (autorità linguistiche, Ammon 2004²): è pertanto di natura extralinguistica. Potrebbe quindi dimostrarsi più efficace rendere conto della presenza di diverse – o delle stesse – strategie nelle varietà standard e non-standard avvalendosi degli strumenti della sociolinguistica.

3. ... alla prospettiva sociolinguistica

L'approccio sociolinguistico ai risultati dell'analisi tipologica richiede innanzitutto un cambio di prospettiva: le strategie di relativizzazione attestate in ciascuna lingua vengono considerate come variabili sociolinguistiche. Sulla base dei giudizi espressi in grammatiche di riferimento e studi linguistici, a ciascuna variabile può essere assegnato uno status sociolinguistico – standard, non-standard o standard/non-standard, qualora i giudizi divergano. Il passo successivo consiste nel rintracciare regolarità, principi o criteri in grado di spiegare la distribuzione delle strategie tra varietà standard e non-standard così com'è emersa dallo studio tipologico. In altre parole, si tratta di indagare secondo quali criteri una strategia di relativizzazione in una determinata lingua riceve l'etichetta 'standard' o 'non-standard', viene cioè considerata sociolinguisticamente neutra o marcata. Questo ci riconduce alla questione più generale della modalità di selezione dei tratti linguistici (fonetico-fonologici, morfosintattici, lessicali, ecc.) che ottengono cittadinanza nella varietà standard di una lingua. Stein (1997) individua tre principi in azione durante il processo di codificazione e mantenimento dello standard:

1. limitazione della variazione interna: solo una parte delle variabili attestate nella lingua (intesa come insieme di varietà) viene inclusa nella varietà standard;
2. imitazione di un modello prestigioso, solitamente esoglossico: nel caso delle lingue europee analizzate si tratta, a seconda del periodo di codificazione dello standard, o delle antiche lingue di cultura (la-

- tino, greco antico, antico slavo ecclesiastico) o di lingue moderne considerate particolarmente prestigiose, quali francese e tedesco;
3. antioralità: lo standard nasce come varietà da impiegare in pratiche testuali caratteristiche della *konzeptionelle Schriftlichkeit* (nel senso di Koch/Oesterreicher 1985) per redigere testi che contengano in sé stessi tutti gli elementi per essere interpretati, evitando per quanto possibile il ricorso a inferenze contestuali.

Da questi principi Stein (1997) deriva quattro criteri corollari che hanno guidato le istanze codificatrici nella scelta delle variabili da includere nello standard:

1. esplicitezza, che deriva dal principio dell'antioralità: ogni significato – lessicale, morfosintattico, pragmatico – deve essere espresso tramite una forma linguistica, per limitare il riferimento al contesto comunicativo. Questo non implica un rapporto uno a uno tra forma e significato, secondo quanto espresso dal principio (funzionale) di iconicità: l'importante è che ogni significato sia espresso tramite una forma; più significati possono essere codificati tramite la stessa forma (v. sotto).
2. compattezza: sono preferibilmente da includere nella varietà standard variabili che condensano significati distinti in una sola forma (ad esempio, il pronome relativo); sono invece da evitare forme polisemiche, che condensano significati diversi a seconda del contesto in una sola forma (ad esempio, congiunzioni polivalenti come il *che* italiano). Parimenti da escludere sono forme cui non sia possibile attribuire un significato preciso (per esempio particelle discorsive o 'parole riempitivo', considerate tipiche dell'oralità).
3. non-ridondanza: una doppia codifica di significati tramite forme diverse è da evitare; in questo modo si ottiene anche un'espressione più compatta. Questo non deve però andare a scapito dell'esplicitezza, che nell'ordine dei criteri occupa il rango più alto.
4. purismo: tratti ritenuti non-nativi perché mutuati da altre lingue, per esempio attraverso fenomeni di contatto linguistico, sono da rigettare. Fanno eccezione, sulla scorta del secondo principio enunciato sopra, forme e strutture presenti nelle lingue considerate modelli esoglossici, che si prestano invece a essere replicate nella varietà standard.

Nel processo di formazione dello standard, le forme che secondo le istanze codificatrici si conformano ai principi e criteri sopra esposti hanno maggiori possibilità di essere accolte nella varietà standard. Si assiste dunque a un processo di giudizio e selezione che conduce da un lato alla (auspicata) riduzione di forme nella varietà standard rispetto a quelle presenti nel diasistema della lingua, dall'altro – all'opposto – a una elaborazione di nuove forme, che possono essere anche replicate su modelli presenti nello standard esoglossico⁶. Il risultato del processo di standardizzazione è una varietà che almeno nelle intenzioni dei codificatori ha un carattere logico e razionale e gode di prestigio in quanto plasmata su un modello esoglossico a sua volta prestigioso – ma che, come nota Weiß (2004), non è per forza naturale e coerente al suo interno.

4. *Una scala di standardità per le strategie di relativizzazione*

Applicando quanto esposto nel par. 3 alle strategie di relativizzazione documentate nelle lingue europee, si può andare in cerca di regolarità nella classificazione delle strategie di relativizzazione come variabili standard o non-standard. Restrungendo l'analisi a un numero limitato di strategie, ossia a quelle documentate in un numero significativo di lingue del campione (e segnatamente alle posizioni 1-7 nella Tabella 2), Murelli (2011: 273-275) ne individua il grado di standardità sulla base dei giudizi contenuti in grammatiche e studi di riferimento e le dispone su una scala partendo da quella considerata più spesso come standard. La scala di standardità così ottenuta ha questa forma:

RPRO, SRE > RPAR (per posizioni alte della Gerarchia di Accessibilità) > RPRO+RE > RPAR+RE > Ø, Ø+RE > RPAR (per posizioni basse della Gerarchia di Accessibilità)

La scala va letta come segue: se una strategia è considerata standard in una lingua, anche tutte quelle alla sua sinistra – se presenti nella lingua – lo saranno. Se osserviamo i criteri che ogni strategia soddisfa (Tabella 3), notiamo una decrescita da sinistra verso destra, con un'ec-

⁶ Questo scenario, parzialmente idealizzato, presuppone che tutte le costruzioni relative siano già esistenti al momento della codifica dello standard. Non potendo qui per motivi di spazio approfondire il rapporto tra sociolinguistica e diacronia, si rimanda a Murelli (2011: 346-367).

cezione costituita dalla strategia del pronome relativo con elemento di ripresa. Benché essa sia conforme a un solo criterio, è collocata più a sinistra di altre strategie che ne soddisfano due. Da un lato questo si può spiegare ricordando la prevalenza del principio di esplicitezza su quello di compattezza, formulata nel par. 3: se esplicita, una strategia è considerata più standard di un'altra strategia compatta, ma non esplicita. Dall'altro lato, pur essendo ridondante, questa strategia contiene un pronome relativo, che, a differenza della strategia immediatamente seguente, 'particella relativa con elemento di ripresa', gode generalmente di un maggiore prestigio.

Tabella 3 - I criteri soddisfatti dalle strategie di relativizzazione presenti sulla scala di standardità

<i>Proprietà</i>	<i>Strategie</i>	<i>RPRO,</i> <i>SRE</i>	<i>RPAR</i> <i>(alte)</i>	<i>RPRO +</i> <i>RE</i>	<i>RPAR +</i> <i>RE</i>	<i>∅ / ∅ +</i> <i>RE</i>	<i>RPAR</i> <i>(basse)</i>
esplicita		+	+/-	+	+	-	-
compatta		+	+/-	-	-	+	+/-
non-ridondante		+	+	-	+	+	+
presente in m. esog.		+	-	-	-	-	-

Per mostrare la validità della scala a livello interlinguistico, Murelli (2011: 276) distribuisce le lingue del campione sulla base delle strategie che in ciascuna di esse sono considerate standard. Ne risulta la Tabella 4, che verrà commentata di seguito, esaminando prima le colonne, quindi il grado di standardità di ciascuna strategia, poi le righe, ossia il "grado di apertura" delle singole lingue nei confronti delle strategie di relativizzazione che costituiscono la scala.

Soffermandosi sulle colonne della Tabella 4, emerge prima di tutto che la strategia del pronome relativo e dell'elemento relativo specializzato, dove esistono, sono considerate standard. Questo non stupisce: come si evince dalla Tabella 3, entrambe rispondono pienamente ai criteri individuati da Stein (1997): sono esplicite, veicolano cioè tutte le proprietà che possono essere codificate da un elemento relativo (o il maggior numero di esse, cfr. nota 4); sono compatte, in quanto costituite da una sola unità morfosintattica; non sono ridondanti, poiché ogni proprietà dell'elemento relativo è codificata una sola volta; sono presenti nei modelli esoglossici di riferimento.

Procedendo lungo la scala, la particella relativa per posizione alte della Gerarchia di Accessibilità – soggetto e oggetto diretto – è considerata standard in tre quarti delle lingue del campione: pur non soddisfacendo tutti i criteri, può essere vista come non ridondante, compatta (se non è polivalente) e (parzialmente) esplicita, proprio perché il suo uso, dove ammesso, è limitato alla relativizzazione delle due posizioni per le quali il ruolo del componente relativizzato è più facilmente ricostruibile.

Tabella 4 - *La disposizione delle lingue del campione sulla scala di standardità*

<i>Lingua</i>	<i>RPRO, SRE</i>	<i>RPAR (alte)</i>	<i>RPRO + RE</i>	<i>RPAR + RE</i>	<i>Ø / Ø + RE</i>	<i>RPAR (basse)</i>
EST, HUN	st.	–	–	–	–	–
FIN, LIT, LAV	st.	nst.	–	–	–	–
TUR	–	nst.	–	nst.	–	–
DUT, GER	st.	nst.	–	nst.	–	–
FRE	st.	nst.	nst.	nst.	–	–
RUS	st.	st.	–	nst.	nst.	nst.
ITA, POR	st.	st.	nst.	nst.	–	nst.
SPA	st.	st.	st.	nst.	–	nst.
ALB, BUL, RUM	st.	st.	st.	st./nst.	–	nst.
BEL, POL, HSB	st.	st.	–	st./nst.	–	–
CZE	st.	st.	–	st./nst.	nst.	–
CAT	st.	st.	st.	st./nst.	–	nst.
DSB	st.	st.	–	st./nst.	–	nst.
BCS	st.	st.	–	st./nst.	–	nst.
ENG	st.	st.	nst.	st./nst.	st./nst.	–
DAN, NOR, SWE	st. (!)	st.	–	st.	st.	–
SLK, SLV, UKR	st.	st.	–	st.	–	–
MAC	st.	st.	st.	st.	–	–
GRE	st. (!)	st.	st.	st.	–	st.
ICE, MLT, GLE	–	st.	–	st.	–	–
BAQ	–	st.	–	st.	–	st.

La combinazione di un pronome relativo con un elemento di ripresa, pur soddisfacendo solo il criterio dell'esplicitezza, è considerata standard in quasi tutte le lingue in cui è attestata⁷. Questo fatto si può forse spiegare considerando che detta strategia è vista come standard

⁷ Nella Tabella 4 è da segnalare, relativamente a questa strategia, la posizione particolare dell'inglese: Radford (2019: cap. 2) riporta diversi esempi, considerandoli

nelle lingue in cui il raddoppiamento clitico dell'oggetto diretto e/o indiretto è a sua volta attestato nella varietà standard – quindi, ad esempio, in spagnolo, ma non in italiano.

La strategia sul quarto gradino della scala, la particella relativa combinata con un elemento di ripresa, è un caso interessante. Le lingue del campione si dividono in tre gruppi sostanzialmente della stessa entità: in dodici lingue la strategia è standard, in otto non-standard, in undici i giudizi di standardità divergono. Se consideriamo i criteri di Stein (1997), notiamo che la strategia ne soddisfa due su quattro: è esplicita e non ridondante, ma non è compatta e, soprattutto, non è presente in nessuno dei modelli esoglossici di riferimento. Si potrebbe ipotizzare che a una maggiore influenza del modello esoglossico nella fase di codificazione dello standard sia corrisposta una maggiore resistenza all'inclusione di questa strategia nella varietà standard, fatto che potrebbe valere particolarmente per lingue di più "antica" codificazione (olandese, tedesco, italiano, francese, portoghese, spagnolo, russo); il contrario sarebbe invece vero per lingue di (ri-)codificazione più recente (sloveno, slovacco, ucraino). In un terzo delle lingue del campione non c'è chiarezza sullo status sociolinguistico di questa strategia: in alcune di esse, per esempio in bosniaco/croato/serbo, le grammatiche di riferimento contengono pareri discordanti sulla sua inclusione nello standard (Kordić 1999: 33-37), ciò che potrebbe far supporre che si tratti di una variabile sociolinguistica "in movimento" o "in evoluzione". In altre lingue, invece, l'etichetta 'standard/non-standard' è dovuta al fatto che sono presenti diverse strategie tipologicamente classificabili come 'particella relativa più elemento di ripresa' caratterizzate da uno status sociolinguistico diverso. È ad esempio il caso dell'inglese: una struttura con elemento di ripresa pronominale, come *the man [that I saw **him**] ...*, attestata a livello dialettale (Herrmann 2005: 70-72), è considerata non-standard, mentre una struttura come *the man [that I'm speaking **of**] ...* – in cui la preposizione può essere interpretata come elemento di ripresa

non-standard (*Supermarkets are now making a big thing about selling wonky vegetables, [which years ago they would just have been discarded]; We need players [who we can count on them in a crisis]*, Radford 2019: 75). L'inglese violerebbe quindi parzialmente la relazione implicazionale alla base della scala di standardità, nel senso che strategie su posizioni più basse del pronome relativo con elemento di ripresa sono considerate, perlomeno in parte, ancora standard.

poiché, pur non essendo un elemento pronominale, codifica il ruolo sintattico relativizzato di obliquo – è considerata standard.

La marca zero (eventualmente accompagnata da un elemento di ripresa) è standard in inglese e nelle lingue scandinave, mentre l'uso di una particella relativa per le posizioni basse della Gerarchia di Accessibilità, dove attestata, è quasi unanimemente classificata come non-standard: pur non essendo ridondante, non è esplicita; se polivalente, non è compatta; infine, non solo non è presente in nessun modello esoglossico, ma è giudicata marcatamente orale, andando così a contrariare con il principio di antioralità.

Prendendo ora in esame le righe della Tabella 4, si può constatare che nelle lingue collocate agli estremi superiore e inferiore è attestato un numero ristretto di strategie: in estone e ungherese solo il pronome relativo, in basco, irlandese, islandese e maltese solo la particella relativa, eventualmente accompagnata da un elemento di ripresa. Queste strategie sono state incluse nello standard di queste lingue poiché erano le uniche attestate. In questo gruppo si potrebbero includere anche le lingue scandinave, che compaiono nella parte bassa della Tabella 4, poco sopra quelle appena citate: qui la strategia del pronome relativo è entrata nella lingua e nello standard tramite traduzioni di testi (religiosi) redatti in latino; essa però non si è imposta sulle strategie più diffuse (particella relativa per posizioni alte della Gerarchia di Accessibilità e particella relativa con elemento di ripresa) ed è rimasta confinata ai registri più formali (ciò che nella tabella è indicato con il simbolo '(!)'), probabilmente perché non c'è stato un forte orientamento verso un modello esoglossico nella fase della codificazione dello standard (Haarmann 1993: 174-175). Una situazione diametralmente opposta è documentabile in francese, olandese e tedesco, lingue in cui, nonostante la presenza di diverse altre strategie, solo quella del pronome relativo è classificata come standard⁸.

⁸ Notiamo che queste lingue fanno parte del cosiddetto *Charlemagne Sprachbund* (van der Auwera 1998), il cuore dello *Standard Average European*, nel quale si riscontra il maggior numero di tratti caratteristici comuni. All'opposto, le lingue in cui il pronome relativo è assente, poco utilizzato o in competizione con diverse altre strategie nella corrispondente varietà standard si trovano ai margini dell'area linguistica europea (lingue germaniche settentrionali, irlandese, basco, macedone, bulgaro, greco...). Cfr. a questo proposito il par. 5.2.

Nella parte centrale della Tabella 4 trovano posto lingue – romanze, slave e balcaniche – la cui varietà standard tollera una variazione interna più spiccata: delle quattro-cinque strategie attestate, la metà sono considerate standard. È possibile supporre che si sia voluto limitare la variazione all'interno della varietà standard, secondo il primo principio enunciato da Stein, privilegiando, nella scelta, soprattutto le strutture presenti nel modello esoglossico di riferimento – quindi la strategia del pronome relativo. Accanto a questa sono attestate la particella relativa per posizioni alte della Gerarchia di Accessibilità, in cui il ruolo del componente relativizzato è più facilmente ricostruibile, e il pronome relativo accompagnato da un elemento di ripresa: pur essendo chiaramente ridondante, questa strategia è preferita alla combinazione di una particella relativa con un elemento di ripresa.

Nella parte medio-bassa della Tabella 4 spicca un gruppo di lingue la cui varietà standard mostra una notevole apertura verso le posizioni più a destra della scala di standardità: macedone, slovacco, sloveno e ucraino. È possibile ravvisare due fattori che hanno probabilmente influito sulle scelte dei codificatori: da un lato la varietà standard di queste lingue si è costituita in tempi recenti, nel tardo XIX secolo o nel corso del XX secolo, di preferenza sulla base di una *koiné* dialettale e senza il condizionamento di un modello esoglossico; dall'altro essa è nata come *Abstand-* e *Ausbauvarietät* (Kloss 1967) in opposizione a uno standard – al più tardi dal momento della codificazione in poi rigettato come varietà esoglossica – che aveva funto da lingua tetto per le varietà locali in questione: lo standard macedone si sviluppa in opposizione al bulgaro, lo slovacco in opposizione al ceco e lo sloveno in opposizione al serbocroato. Forse anche per rimarcare la distanza (*Abstand*) rispetto alla lingua tetto, i codificatori hanno ammesso nella nuova varietà strategie che nelle rispettive lingue tetto, pur se attestate, non erano considerate standard (Murelli 2011: 270). Ne risulta che le varietà standard di queste lingue sono più aperte verso le posizioni basse della scala di standardità.

Un caso estremo è costituito del greco. Fino al 1976 la varietà standard era la *katharévusa*, modellata sul greco antico, in cui solo la strategia del pronome relativo valeva come standard; quando la *dhimotiki*, basata su una *koiné* dialettale, l'ha sostituita, tutte le strategie basate sulla particella relativa – eventualmente accompagnata da un elemento di ripresa – hanno cambiato status sociolinguistico, diven-

tando standard, mentre la strategia del pronome relativo, pur rimanendo standard, è stata confinata ai registri più formali (Haarmann 1993: 179, Murelli 2011: 366).

5. Sulla validità della scala di standardità

In questa sezione si esamineranno tre questioni concernenti la scala di standardità formulata nel par. 4: nel par. 5.1 si evidenzieranno punti di forza e punti deboli della scala; nel par. 5.2 si indagherà il rapporto tra la scala e il concetto di *Standard Average European*; nel par. 5.3 si sonderà la capacità predittiva della scala nel quadro della formazione di varietà cosiddette neo-standard nelle lingue europee.

5.1 La scala di standardità: punti di forza e punti critici

Un primo ambito di riflessione concerne la legittimità stessa della scala di standardità. Ammon (2004²: 278) prende le distanze dalla formulazione di simili scale: “it seems possible, in principle, to develop scales of standardness [...] though their general validity may remain questionable.” Questo *caveat* assume un rilievo anche maggiore nel caso di una scala di standardità che aspiri a essere valida a livello interlinguistico. A livello *intralinguistico*, infatti, appare plausibile individuare variabili sociolinguistiche – fonetiche, morfologiche, sintattiche, ecc. – e valutarne il grado di standardità tramite strumenti *ad hoc*: vuoi somministrando interviste e questionari a informanti, vuoi ricercando le variabili all’interno di corpora, vuoi incrociando dati raccolti sul campo con quanto riportato in grammatiche di riferimento. A livello *interlinguistico* la situazione si complica, specialmente se le lingue coinvolte sono numerose, come è il caso del presente studio areale. Due sono le critiche che si possono muovere alla scala formulata nel par. 4. Da un lato, non tutte le strategie emerse dall’analisi tipologica sono rappresentate sulla scala⁹: di conseguenza il quadro tipologico non è riprodotto nella sua totalità a livello sociolinguistico. Dall’altro, la granularità della scala è piuttosto grossa: all’interno di ciascuna lingua sono spesso presenti più istanze di una stessa strategia (più di un pronome relativo o di una particella relativa, v. Murelli 2011: 193ss.,

⁹ Questo è dovuto solitamente a mancanza di informazioni circa lo status di alcune strategie documentate in un numero esiguo di lingue.

278ss.) i cui contesti d'uso mostrano una distribuzione più articolata rispetto alla mera opposizione 'standard vs. non-standard'.

Va inoltre aggiunto che, mentre le strategie di relativizzazione sono state classificate sulla base di tre parametri, la scala è basata esclusivamente sul parametro 'elemento relativo'; solo in parte viene considerato il parametro 'posizioni relativizzate', con la distinzione tra particella relativa per posizioni alte e posizioni basse della Gerarchia di Accessibilità. Una più approfondita analisi di quest'ultimo parametro in termini sociolinguistici permetterebbe di operare distinzioni più fini in seno alle singole strategie. Due esempi:

1. il pronome relativo combinato con un elemento di ripresa, dove è considerato standard, lo è per le posizioni di oggetto diretto e indiretto – non quindi per posizioni più basse – e solo se l'elemento di ripresa è un clitico¹⁰;
2. la particella relativa con elemento di ripresa può essere soggetta a restrizioni d'uso in lingue in cui è etichettata come standard: per esempio, in sloveno è standard per le posizioni di oggetto diretto e indiretto, mentre per posizioni più basse lo standard ammette solo il pronome relativo (Murelli 2011: 229-230)¹¹.

Pur nella consapevolezza che la scala di standardità presenta diversi punti deboli, occorre sottolineare quello che a parere di chi scrive è il suo vero punto di forza: riportare trentasei lingue europee su un'unica scala fornisce un quadro globale – seppur, come detto, a grana grossa – della distribuzione sociolinguistica di strutture fin qui studiate dal punto di vista tipologico esclusivamente a livello di varietà standard, escludendo cioè la dimensione sociale della diversità linguistica. Il cambio di prospettiva contestuale alla formulazione della scala di standardità ha il merito di portare in primo piano la complessa architettura delle lingue analizzate: ciascuna di esse è un diasistema costituito da varietà all'interno delle quali possono comparire diverse strategie di relativizzazione; nessuna di queste può essere esclusa quando

¹⁰ Per ulteriori differenziazioni nelle lingue in cui è attestata questa strategia cfr. Murelli (2011: 271-272).

¹¹ Casi come questo illustrano le inconsistenze a volte presenti nelle varietà standard, al cui confronto le varietà non-standard risultano più regolari o coerenti: nelle varietà non-standard di sloveno la strategia in questione può essere applicata anche alle posizioni più basse della Gerarchia di Accessibilità.

ci si cimenta con un'analisi tipologica. La scala e la distribuzione delle lingue su di essa (Tabella 4) mostrano a colpo d'occhio che strategie di relativizzazione tipologicamente simili possono presentare uno status sociolinguistico diverso in lingue diverse e che esiste uno strumento – la scala, appunto – in grado di catturare e riprodurre, almeno in parte, questo tipo di variazione. Un ideale completamento della scala sarebbe la possibilità di “zoomare” su ciascuna lingua o strategia per evidenziare peculiarità a livello di singole lingue (esistenza di diverse strategie tipologicamente identiche che hanno uno status sociolinguistico differente; restrizioni d'uso di determinate strategie, ecc.).

5.2 Scala di standardità e Standard Average European

Un secondo ambito di riflessione riguarda il rapporto tra la posizione delle lingue sulla scala e lo *Standard Average European*. La Tabella 4 potrebbe essere interpretata come una conferma dell'esistenza, nell'area linguistica europea, di uno *Standard Average European* composto da lingue che condividono tratti (morfosintattici) comuni. Secondo Haspelmath (2001: 1494s.) uno di questi tratti è la formazione di frasi relative, che avverrebbe preminentemente tramite la strategia del pronome relativo e in seconda battuta tramite quella della particella relativa. Questo è vero per tutte le lingue del campione; l'indagine sociolinguistica condotta nei paragrafi 3 e 4 permette però di precisare questa considerazione. Il pronome relativo è presente e considerato standard in tutte le lingue a eccezione di un gruppo situato ai margini dell'area linguistica europea (da nord a sud: islandese, irlandese, basco, maltese), in cui è assente, ed è confinato ai registri più formali della varietà standard in un altro gruppo di lingue, collocate anch'esse ai margini dell'area (lingue scandinave a nord, greco a sud). In queste lingue la strategia usata di *default* è quella della particella relativa, eventualmente accompagnata da un elemento di ripresa. Viceversa, un gruppo di lingue considera standard solo la strategia del pronome relativo, a prescindere dalla presenza di altre strategie: si tratta da un lato di lingue situate al centro dello *Standard Average European*, nel cosiddetto *Charlemagne Sprachbund* (van der Auwera 1998), dall'altro delle lingue baltiche e ugro-finniche – queste ultime disposte in posizione più periferica, verso nordest.

Alla luce di quanto detto, possiamo concludere che la metà sinistra della Tabella 4 conferma quanto sostenuto da Haspelmath (2001) a

proposito delle frasi relative nello *Standard Average European*; di converso, appare lecito chiedersi insieme a Seiler (2019) se esistano tratti caratteristici di un *Non-Standard Average European* e, in caso affermativo, quali siano. Spostando l'attenzione verso la metà destra della Tabella 4, osserviamo che alcune delle strategie considerate più spesso non-standard – particella relativa (applicata alle posizioni basse della Gerarchia di Accessibilità) e elemento zero – non codificano esplicitamente il ruolo del componente relativizzato nella frase relativa. Alla destra di queste ultime sarebbero collocabili altre strategie che non sono state incluse nella scala di standardità e nella Tabella 4 poiché attestata in un numero esiguo di lingue: si tratta di elementi combinati composti da due elementi semplici (pronome relativo o elemento relativo specializzato più particella relativa) o da un elemento semplice e una particella non relativa (complementatore o altra particella, solitamente di origine deittica, cfr. Murelli 2011: 101ss.), salvo poche eccezioni classificati come non-standard nelle lingue in cui sono attestati. Risalendo la Tabella 4 verso sinistra troviamo la strategia della particella relativa con pronomi di ripresa, di cui si è discusso nel par. 4. Se osserviamo la sua distribuzione nelle lingue del campione, notiamo che è parallela a quella del pronome relativo, ma di segno opposto: nelle lingue costituenti lo *Charlemagne Sprachbund* è non-standard, mentre alla periferia dell'area linguistica europea (islandese, lingue scandinave, irlandese, basco, macedone, greco) viene più di frequente considerata standard.

Tornando all'interrogativo di Seiler (2019), pare possibile affermare che, mentre lo *Standard Average European* mostra una spiccata preferenza per la strategia del pronome relativo (e solo in seconda battuta e non in tutte le lingue per la particella relativa per le posizioni alte della Gerarchia di Accessibilità), il *Non-Standard Average European* converge sulla strategia della particella relativa accompagnata da un elemento di ripresa – attestata nell'86% delle lingue del campione – e, in misura minore, su quella della particella senza elemento di ripresa per le posizioni basse della Gerarchia di Accessibilità, attestata solo in un terzo delle lingue¹². Accanto a queste esistono altre strategie, introdotte da elementi combinati, attestate in un numero troppo esiguo di lingue perché si possa parlare di convergenza; queste sono state esclu-

¹² Benché documentata solo in 12 su 36 lingue, non è escluso che ulteriori e più approfondite analisi ne rivelino l'esistenza in altre.

se in fase di codificazione dalla corrispondente varietà standard perché non si conformano ai criteri enunciati nel par. 3 – nella fattispecie violano il criterio di compattezza o di non-ridondanza.

Il fatto che attorno al polo del non-standard si concentrino strategie che non codificano il ruolo sintattico del componente relativizzato o che lo codificano tramite un elemento di ripresa rivela un'affinità tra il *Non-Standard Average European* e la distribuzione tipologica delle strategie di relativizzazione nelle lingue del mondo. I risultati dell'indagine di Comrie & Kuteva (2013) indicano che la strategia del *gap* (ossia una strategia che non codifica il ruolo sintattico del componente relativizzato nella frase relativa) è la più largamente diffusa a livello globale per relativizzare la posizione di soggetto (125 lingue su 166) – ma non in Europa¹³, dove invece domina il tipo del pronome relativo (10 su 15 lingue). Per relativizzare la posizione di obliquo, invece, la strategia del *gap* concorre con quella della *pronoun retention*, in cui il ruolo sintattico del componente relativizzato è codificato tramite un pronome di ripresa (rispettivamente 55 e 20 lingue su 112 a livello globale); in Europa domina di nuovo il pronome relativo (11 su 15 lingue). La distribuzione delle strategie nel *Non-Standard Average European* si rivela dunque più simile a quella globale: esso mostra, sotto questo aspetto, tratti meno idiosincratici dello *Standard Average European* e – dato interessante – si conforma maggiormente alla distribuzione delle strategie di relativizzazione presenti nelle varietà standard (!) delle lingue poste ai margini dell'area linguistica europea.

5.3 La scala di standardità e le varietà neo-standard

Un ultimo ambito di riflessione concerne la capacità predittiva della scala di standardità nel quadro dello sviluppo, in alcune lingue europee, di varietà cosiddette neo-standard, ossia di varietà nate tramite un processo di demotizzazione delle varietà standard tradizionali. Queste ultime erano nate per essere impiegate in pratiche testuali ascrivibili alla *konzeptionelle Schriftlichkeit*; la loro diffusione e padronanza era limitata a una parte ristretta della popolazione. Specialmente nella seconda metà del XX secolo si assiste però a una “popolarizzazione” della varietà standard, che diventa la prima lingua appresa da buona

¹³ O, per meglio dire, nelle varietà standard delle lingue europee considerate da Comrie & Kuteva (2013).

parte dei parlanti, sostituendo in questo ruolo varietà basilettali e declinazioni regionali dello standard stesso. Uno degli effetti di questa popolarizzazione è proprio la formazione di varietà neo-standard (Auer 2018: 38-40).

Standard tradizionali e neo-standard condividono tratti simili: entrambi sono sovraregionali – anche eventuali tratti originariamente regionali, entrando nel neo-standard, perdono il loro carattere diotopicamente marcato – e godono di alto prestigio. Mentre lo standard tradizionale basa questo prestigio su caratteristiche quali formalità, letterarietà, impersonalità e oggettività, il neo-standard lo deriva da qualità in parte opposte – informalità, modernità, oralità, soggettività –, che ne fanno la varietà di riferimento ideale per l'utilizzo in una società globalizzata in cui *mass media* e reti sociali assumono un ruolo sempre più dominante (Auer 2017: 371-373, Auer 2018: 44-51)¹⁴. Il neo-standard è nato come varietà orale, affermandosi poi nello scritto ed estendendo il suo uso da pratiche testuali informali (e-mail, messaggi, post), legate alla *Sprache der Nähe*, a pratiche più formali (p. es. testi giornalistici, Auer 2018: 42), in cui è tradizionalmente impiegata la *Sprache der Distanz* (nel senso di Koch & Oesterreicher 1985).

Una varietà standard si situa ed evolve all'interno di un "campo di forza" costituito da quattro componenti (Ammon 2004²: 276-277): codificatori, autorità normative, parlanti/scriventi modello, esperti (linguisti). L'equilibrio e i mutui rapporti tra queste autorità linguistiche permettono lo sviluppo e il mantenimento di uno standard che assolva ai compiti per cui è stato codificato, che possa cioè ad esempio essere insegnato nelle scuole e utilizzato nella comunicazione quotidiana in determinate pratiche testuali.

Ora, ci si può chiedere che cosa succede se le istanze che costituiscono il "campo di forza" cambiano (Ballarè 2020: 474-475). È il caso della formazione di varietà neo-standard: non è più il canone letterario a costituire il corpus di testi di riferimento, né gli estensori di tali testi vengono più considerati parlanti/scriventi modello, sostituiti da esponenti del mondo politico e massmediatico (giornalisti della carta stampata, di radio, televisione e internet), e più in generale da

¹⁴ Occorre precisare con Auer (2018: 52) che il neo-standard non sta (ancora) scalzando lo standard tradizionale, il cui prestigio originario è rimasto inalterato; sono cambiate piuttosto le pratiche testuali e le situazioni comunicative in cui quest'ultimo è utilizzato, che si concentrano soprattutto in ambiti di formalità alta o medio-alta.

“cultured, well educated speakers and writers” (Berruto 2017: 36). Questo può portare a una ridefinizione di quali strutture linguistiche siano accettabili all’interno di una varietà considerata standard. Nel caso del neo-standard, come sottolinea Auer (2017: 371), l’estensione del dominio sociolinguistico di applicazione di questa varietà, che comprende un ampio spettro di pratiche testuali, richiede, per poter funzionare, una maggiore flessibilità strutturale: una delle possibilità per ottenerla è incrementare la variazione interna, per esempio includendo variabili sociolinguistiche che secondo i canoni dello standard tradizionale verrebbero classificate come non-standard.

Come si ripercuote l’incrementata variazione interna sul fenomeno oggetto di questo contributo, e in particolare sulla scala di standardità formulata nel par. 4? Un’indagine estesa a tutte le lingue del campione non potrà che essere oggetto di studi futuri; di seguito ci si limiterà a indicare alcune tendenze riscontrabili in italiano e in tedesco.

Nell’italiano neo-standard si osserva un’accreciuta tolleranza verso la strategia del pronome relativo accompagnato da un elemento di ripresa, come in (7).

(7) italiano

Poco prima di commettere il folle gesto ha deciso di rispondere alla telefonata del fidanzato [a cui gli ha detto]: “Mi vedi, alza la testa” [...] (Voce di Napoli, 10/04/2018)

Cerruti (2017: 78-79) nota che essa è documentata anche in testi scritti di carattere formale, tra cui articoli giornalistici. Come accennato sopra, i giornalisti figurano tra i parlanti/scriventi modello per le varietà neo-standard: il fatto che impieghino con una certa regolarità una costruzione tradizionalmente considerata non-standard in testi situabili vicino al polo della *konzeptionelle Schriftlichkeit* può essere interpretato come un segnale che la costruzione in questione sta acquisendo o ha acquisito cittadinanza nel neo-standard. Una breve ricerca condotta in Murelli (submitted) mostra occorrenze di questa strategia in particolare in testate di stampa locale, meno in quelle nazionali; un’indagine più approfondita potrà verificarne la consistenza di questa tendenza.

Nel tedesco neo-standard si profilano due fenomeni paralleli: il pronome *was* e i gli avverbi costituiti da *wo*+preposizione, nello standard tradizionale impiegabili esclusivamente per relativizzare rispettivamente le posizioni soggetto/oggetto diretto e obliquo e applicabili

a un ristretto gruppo di antecedenti di genere neutro (Duden 2009: 1031-1032) – tra cui il dimostrativo *das*, pronomi indefiniti (*etwas*, *nichts* ‘qualcosa, niente’), aggettivi sostantivati (*das Gute* ‘la cosa buona’, *das Beste* ‘il meglio’, *das Erste* ‘la prima cosa’) – paiono estendere il proprio campo d’impiego anche ad antecedenti sostantivali, neutri nel caso di *was*, dei tre generi nel caso di *wo*+preposizione, come illustrato negli esempi seguenti. In (8) l’elemento relativo *was* riprende il sostantivo neutro *das Geld*; in (9) e (10) gli antecedenti maschile e femminile *der Schlüssel* und *[ei]ne Sache* sono relativizzati da *wo*+preposizione – in (10) interviene anche il *preposition stranding*¹⁵.

- (8) tedesco
Denn das Geld, [was Apple
 perché DET denaro RPRO.ACC.SG.N Apple
damit verdienen würde,] läge weiter weit über
 con.ciò guadagnerebbe starebbe ancora lontano sopra
dem der Konkurrenz.
 quello DET concorrenza
 ‘Perché il denaro che Apple ci guadagnerebbe supererebbe ancora di gran lunga quello della concorrenza.’ (Börse am Sonntag, 12/10/2018)

- (9) tedesco
In einem Wahlbüro in Kinshasa
 in DET ufficio.elettorale a Kinshasa
fehlte der USB-Stick mit dem digitalen
 mancava DET chiavetta.USB con DET digitale
Schlüssel, [womit sich die Wahlmaschinen starten lassen.]
 chiave RPAR.CON REFL DET urne.elettroniche azionare fanno
 ‘In un ufficio elettorale di Kinshasa mancava la chiavetta USB con il codice che serve ad azionare le urne elettroniche.’ (taz, 31/12/2018)

¹⁵ Come si vede, negli esempi (9)-(10) il morfema *wo* è glossato ‘RPAR.’ In tedesco gli elementi ‘*wo*+preposizione’ sono chiamati *Relativadverbien*; seguendo questa nomenclatura, rappresenterebbero istanze di elementi relativi specializzati. Il fatto che la preposizione possa essere separata e che al suo posto possa comparire – in varietà non-standard – un elemento di ripresa pronominale ‘*da(r)*+preposizione’ (in tedesco *Pronominaladverb*), come in *Das Thema, [wo wir drüber gesprochen haben]* ‘L’argomento di cui (lett. che ne) abbiamo parlato’, fa pendere per una segmentazione ‘particella relativa con elemento di ripresa’: la parte *wo* codifica il legame tra frase matrice e frase relativa, mentre la preposizione o il *Pronominaladverb* veicolano il ruolo del componente relativizzato.

(10) tedesco

ja Moni das ist ne sache [wo
 sì Moni DEM è DET cosa RPAR
 man immer mit rechnen muss]

IMPERS sempre con contare deve

‘Sì, Moni, questa è una cosa con cui bisogna sempre fare i conti.’

(spin.de)

Entrambi i fenomeni sono già documentati in varietà non-standard – la prima in varietà dialettali (Fleischer 2005: 178, Murelli 2012), la seconda nelle *Umgangssprachen* regionali (Murelli submitted): anche in questo caso si può ipotizzare che il neo-standard abbia incluso al suo interno costruzioni che le grammatiche di riferimento non indicano come standard o sulle quali i giudizi di standardità divergono¹⁶. In effetti anche in questo caso non mancano occorrenze in testi giornalistici – cfr. (9) sopra e ulteriori esempi in Murelli (submitted).

Il tedesco neo-standard presenta dunque una situazione che potremmo definire di transizione: mentre nello standard tradizionale solo la strategia del pronome relativo (*der/die/das* e *welcher/welche/welches*) è ammessa per la relativizzazione di antecedenti nominali, nel neo-standard il pronome *was* può sostituire *das* per relativizzare soggetto e oggetto diretto se l’antecedente è neutro; la combinazione *wo*+preposizione può sostituire il pronome relativo per relativizzare la posizione obliquo se l’antecedente – non importa di quale genere – non è animato. Occorre precisare che la preposizione può essere separata dall’elemento *wo*. Il *preposition stranding* sembra però (ancora) caratterizzare registri meno controllati, come risulta da una prima ricerca effettuata in Murelli (submitted): in occorrenze estratte da testi a stampa la preposizione non è separata da *wo*; il fenomeno ricorre invece in contesti meno sorvegliati, come i forum di discussione (v. (10) sopra).

Ci si può chiedere ora se ci sia una relazione tra l’apertura verso nuove strategie di relativizzazione nel neo-standard e la scala di stan-

¹⁶ Sull’utilizzo di *was* in alternativa a *das* per relativizzare antecedenti nominali si veda il dettagliato studio di Brandt & Fuß (2019). Murelli (submitted) constata che le grammatiche tedesche di riferimento esprimono pareri differenti circa l’utilizzo di *wo*+preposizione per relativizzare antecedenti sostantivali, come in *der Hammer [mit dem / womit ich arbeite]* ‘il martello con cui lavoro’: ora sono considerate semplici alternative alla strategia del pronome relativo, ora se ne sconsiglia l’utilizzo, ora se ne dichiara l’uso come arcaico, sebbene ancora diffuso.

dardità. Per quanto riguarda l'italiano, notiamo che questa apertura segue l'ordine delle strategie sulla scala: il pronome relativo accompagnato da elemento di ripresa è la prima delle strategie considerate non-standard¹⁷. In tedesco il quadro non è così chiaro: come si è detto nel par. 4, lo standard tradizionale mostra poca o nessuna tolleranza verso strategie diverse dal pronome relativo. Tuttavia, dati tratti da testi giornalistici evidenziano una parziale apertura verso la strategia con elemento relativo *wo*+preposizione se l'antecedente è [-animato]. L'elemento *was*, che in alcune varietà dialettali ha perso la sua natura pronominale e funziona come particella relativa, relativizza solo antecedenti neutri nei testi giornalistici analizzati, conservando così le sue caratteristiche morfosintattiche e il suo status di pronome. Come tale, non può essere (ancora) considerato istanza della strategia della particella relativa. Nel caso del tedesco si può quindi parlare di un'apertura "condizionata" verso strategie tradizionalmente considerate non-standard, apertura che non pare seguire la scala di standardità.

6. Conclusioni

In questo contributo si è voluto mostrare come la tipologia linguistica e la sociolinguistica possano interagire fruttuosamente. Per esemplificare come può funzionare questa interazione si è scelto di analizzare le strategie di relativizzazione in trentasei lingue europee. Le strategie attestate nel campione sono state classificate secondo tre parametri ampiamente impiegati negli studi tipologici sulle frasi relative. Si è cercato poi di rendere conto delle differenze tra le strategie presenti nelle varietà standard e in quelle non-standard sulla base di principi funzionali. Poiché questo approccio non si è rivelato efficace, si è at-

¹⁷ Il grado di non-standardità delle strategie escluse dallo standard tradizionale è stato indagato in Aureli (2003), che ha condotto un'indagine tra giovani parlanti romani sottoponendo loro testi scritti e orali contenenti diverse strategie di relativizzazione e chiedendo di esprimere un giudizio di accettabilità. Per quanto riguarda i testi scritti, emerge che la strategia del pronome relativo con elemento di ripresa è stata giudicata più accettabile di quella della particella relativa con elemento di ripresa e quest'ultima più accettabile di quella della particella relativa per posizioni basse della Gerarchia di Accessibilità. È interessante notare come il grado di accettabilità espresso dagli informanti risulti parallelo alla scala di standardità; la prima strategia a essere inclusa nel neo-standard è stata proprio quella considerata più accettabile, ossia meno 'non-standard' dagli informanti consultati da Aureli.

tuato un cambio di prospettiva: le strategie di relativizzazione sono state considerate come variabili sociolinguistiche cui si è attribuito uno status corrispondente – standard o non-standard. Sulla scorta di Stein (1997) sono stati identificati principi e criteri che possono aver guidato le scelte delle istanze codificatrici durante il processo di formazione dello standard: solo le strategie che vi si confacevano sono state incluse nella nuova varietà. L'analisi di quali strategie siano considerate (non-)standard ha permesso di formulare una scala di standardità di tipo implicazionale: il grado di standardità delle strategie che vi sono rappresentate decresce più ci si sposta verso le posizioni basse della scala. Consapevoli del *caveat* espresso da Ammon (2004²) circa l'opportunità di elaborare scale di standardità, nel par. 5 si è esaminata criticamente la scala, individuandone punti di forza e di debolezza e saggiandone la validità in relazione al tema dello *Standard Average European* e della formazione di varietà neo-standard in alcune lingue europee. Ne sono risultati tre ordini di considerazioni.

1. La scala di standardità offre una panoramica della distribuzione sociolinguistica delle strategie di relativizzazione attestate più di frequente nelle lingue europee del campione. Il quadro che ne emerge è inevitabilmente a grana grossa e non tiene conto del fatto che a livello intralinguistico diverse strutture morfosintattiche che realizzano la stessa strategia possono avere uno status sociolinguistico differente. Per rendere conto di queste peculiarità sarebbe necessario declinare la scala a livello intralinguistico per ciascuna delle lingue considerate, cosa che esula dagli scopi di uno studio areale.
2. L'ordine delle strategie sulla scala di standardità rafforza l'ipotesi dell'esistenza di caratteristiche comuni condivise dalle lingue che formano lo *Standard Average European*; dall'altra parte, il tentativo di individuare tratti di un ipotetico *Non-Standard Average European* mostra che questo, dal punto di vista tipologico, si conformerebbe maggiormente a tendenze valide a livello globale.
3. L'apertura verso nuove strategie di relativizzazione nel contesto della formazione di varietà neo-standard può avvenire seguendo l'ordine proposto dalla scala di standardità: questo appare vero per l'italiano, meno invece per il tedesco.

Integrare tipologia e sociolinguistica contribuisce dunque a gettare nuova luce su questioni sensibili alla dimensione sociale della diversità linguistica; in particolare, le scale di standardità possono rivelarsi strumenti utili per modellare a livello intra- e interlinguistico la distribuzione di variabili sociolinguistiche e – ma questa è un’ipotesi che va senza dubbio ulteriormente testata – per rappresentare eventuali cambiamenti in atto, come la riformulazione dello status sociolinguistico di singole variabili.

Elenco delle abbreviazioni utilizzate

3 = terza persona, ACC = accusativo, CL = clitico, COMP = complementatore, DAT = dativo, DEM = dimostrativo, DET = determinante, F = femminile, FM = frase matrice, FR = frase relativa, IMPERS = pronomi impersonali, INSTR = strumentale, M = maschile, NEG = negazione, NOM = nominativo, nst. = non-standard, N = neutro, Ø = marca zero, PL = plurale, RE = elemento di ripresa, REFL = pronomi riflessivi, RPAR = particella relativa, RPRO = pronomi relativi, SG = singolare, SRE = elemento relativo specializzato, st. = standard

Riferimenti bibliografici

- Ammon, Ulrich. 2004². Standard variety. In Ammon, Ulrich & Dittmar, Norbert & Mattheier, Klaus J. & Trudgill, Peter (a cura di), *Sociolinguistics. An international handbook of the science of language and society*, 273-283. Berlin & New York: Mouton de Gruyter.
- Andrews, Avery D. 2007. Relative clauses. In Shopen, Timothy (a cura di), *Language typology and syntactic description*. Vol. 2: *Complex constructions*, 206-236. Cambridge: Cambridge University Press.
- Auer, Peter. 2017. The neo-standard of Italy and elsewhere in Europe. In Cerruti, Massimo & Crocco, Claudia & Marzo, Stefania (a cura di), *Towards a new standard: Theoretical and empirical studies on the restandardization of Italian*, 365-374. Berlin & New York: Mouton de Gruyter.
- Auer, Peter. 2018. The German neo-standard in a European context. In Stichel, Gerhard (a cura di), *National language institutions and national languages. Contributions to the EFNIL Conference 2017 in Mannheim*,

- 37-56. Budapest: Research Institute for Linguistics, Hungarian Academy of Sciences.
- Aureli, Massimo. 2003. Pressione dell'uso sulla norma. Le relative non standard nei giudizi degli utenti. *Studi italiani di linguistica teorica e applicata* 32(1). 45-67.
- van der Auwera, Johan. 1998. Conclusion. In van der Auwera, Johan (a cura di), *Adverbial constructions in the languages of Europe*, 813-832. Berlin & New York: Mouton de Gruyter.
- van der Auwera, Johan. 2011. Standard Average European. In van der Auwera, Johan & Kortmann, Bernd (a cura di), *The languages and linguistics of Europe: A comprehensive guide*, 291-306. Berlin & New York: Mouton de Gruyter.
- Ballarè, Silvia. 2020. L'italiano neo-standard oggi: stato dell'arte. *Italiano LinguaDue* 12(2). 469-492.
- Balliet, Pierre. 1997. *La relative en Alsace Bossue*. Stuttgart: Hans-Dieter Heinz.
- Berruto, Gaetano. 2012 [1987]. *Sociolinguistica dell'italiano contemporaneo*. Roma: Carocci.
- Berruto, Gaetano. 2017. What is changing in Italian today? Phenomena of restandardization in syntax and morphology: an overview. In Cerruti, Massimo & Crocco, Claudia & Marzo, Stefania (a cura di), *Towards a new standard: Theoretical and empirical studies on the restandardization of Italian*, 31-60. Berlin & New York: Mouton de Gruyter.
- Brandt, Patrick & Fuß, Eric. 2019. Relativpronomenselektion und grammatische Variation: *was* vs. *das* in attributiven Relativsätzen. In Fuß, Eric & Konopka, Marek & Wöllstein, Angelika (a cura di), *Grammatik im Korpus*, 91-209. Tübingen: Narr.
- Cerruti, Massimo. 2017. Changes from below, changes from above. Relative constructions in contemporary Italian. In Cerruti, Massimo & Crocco, Claudia & Marzo, Stefania (a cura di), *Towards a new standard: Theoretical and empirical studies on the restandardization of Italian*, 61-88. Berlin & New York: Mouton de Gruyter.
- Comrie, Bernard & Kuteva, Tanja. 2013. Relativization Strategies. In Dryer, Matthew S. & Haspelmath, Martin (a cura di), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology (URL: <http://wals.info/chapter/s8>, consultato il 15/02/2021).
- Cristofaro, Sonia & Giacalone Ramat, Anna. 2007. Relativization strategies in the languages of Europe. In Ramat, Paolo & Roma, Elisa (a cura di),

- Europe and the Mediterranean as linguistic areas: Convergences from a historical and typological perspective*, 63-93. Amsterdam & Philadelphia: John Benjamins.
- Dudenredaktion. 2009⁸. *Duden. Die Grammatik*. Mannheim, Wien & Zürich: Dudenverlag.
- Fleischer, Jürg. 2005. Relativsätze in den Dialekten des Deutschen: Vergleich und Typologie. *Linguistik Online* 24(3). 171-186.
- Glaser, Elvira. 1997. Dialektsyntax: eine Forschungsaufgabe. In *Bericht über das Jahr 1996. Schweizerdeutsches Wörterbuch. Schweizerisches Idiotikon*, 11-30. Zürich (URL: https://www.idiotikon.ch/Texte/Jahresberichte/Id_Jahresbericht_1996.pdf, consultato il 15/02/2021).
- Haarmann, Harald 1993 *Die Sprachenwelt Europas*. Frankfurt & New York: Campus.
- Haspelmath, Martin. 2001. The European linguistic area: Standard Average European. In Haspelmath, Martin & König, Ekkehard & Oesterreicher, Wulf & Raible, Wolfgang (a cura di), *Language typology and language universals. An international handbook*, 1492-1510. Berlin & New York: Mouton de Gruyter.
- Herrmann, Tanja. 2005. Relative clauses in English dialects of the British Isles. In Kortmann, Bernd & Herrmann, Tanja & Pietsch, Lukas & Wagner, Susanne, *A comparative grammar of British English dialects: Agreement, gender, relative clauses*, 21-124. Berlin & New York: Mouton de Gruyter.
- Hualde, José Ignacio. 1992. *Catalan*. London & New York: Routledge.
- Keenan, Edward L. & Comrie, Bernard. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8. 63-99.
- Kloss, Heinz. 1967. Abstand languages and ausbau languages. *Anthropological linguistics* 9(7). 29-41.
- Koch, Peter & Oesterreicher, Wulf. 1985. Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36. 15-43.
- Kordić, Snježana. 1999. *Der Relativsatz im Serbokroatischen*. München: Lincom Europa.
- Kortmann, Bernd (a cura di). 2004. *Dialectology meets typology. Dialect grammar from a cross-linguistic perspective*. Berlin & New York: Mouton de Gruyter.

- Lehmann, Christian. 1984. *Der Relativsatz. Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik*. Tübingen: Narr.
- Murelli, Adriano. 2011. *Relative constructions in European non-standard varieties*. Berlin & New York: Mouton de Gruyter.
- Murelli, Adriano. 2012. Das Geheimnis, *das* oder *was* du mir verraten hast? *Das* oder *was* als Relativpronomen. In Konopka, Marek & Schneider, Roman (a cura di), *Grammatische Stolpersteine digital: Festschrift für Bruno Strecker zum 65. Geburtstag*, 145-152. Mannheim: Institut für Deutsche Sprache.
- Murelli, Adriano (submitted). Relativsätze im Italienischen und im Deutschen. Ein Vergleich unter Berücksichtigung von Nicht-Standard-Varietäten. *Linguistik Online*.
- Petr, Jan (red). 1987. *Mluvnice češtiny* [Grammatica ceca]. Vol. 3: *Skladba* [Sintassi]. Praha: Academia.
- Radford, Andrew. 2019. *Relative Clauses: Structure and Variation in Everyday English*. Cambridge: Cambridge University Press.
- Seiler, Guido. 2019. Non-Standard Average European. In Nievergelt, Andreas & Rübekell, Ludwig (a cura di), *Raum und Sprache: athe in palice, athe in anderu sumeuuelicheru stedi. Festschrift für Elvira Glaser zum 65. Geburtstag*, 541-554. Heidelberg: Winter.
- Stein, Dieter. 1997. Syntax and varieties. In: Cheshire, Jenny & Stein, Dieter (a cura di), *Taming the vernacular. From dialect to written standard language*, 35-50. London & New York: Longman.
- de Vries, Mark. 2002. *The syntax of relativization*. Utrecht: LOT.
- Weiß, Helmut. 2004. A question of relevance. Some remarks on standard languages. *Studies in Language* 28(3). 648-674.
- Zifonun, Gisela. 2001. *Grammatik des Deutschen im europäischen Vergleich: Der Relativsatz*. Mannheim: Institut für Deutsche Sprache.

Autrici e autori

Silvia Ballarè – Dipartimento di Filologia Classica e Italianistica,
Alma Mater Studiorum - Università di Bologna, Bologna, silvia.ballare@unibo.it

Guglielmo Inglese – Department of Linguistics, KU Leuven, Leuven,
Belgium & FWO - Research Foundation Flanders, [guglielmo.inglese@kuleuven.be](mailto:inglese@kuleuven.be)

Massimo Cerruti – Dipartimento di Studi Umanistici, Università di
Torino, Torino, massimosimone.cerruti@unito.it

Nicola Grandi – Dipartimento di Filologia Classica e Italianistica
Alma Mater Studiorum - Università di Bologna, Bologna, nicola.grandi@unibo.it

Marco Angster – Dipartimento di Linguistica, Università di Zara,
Zara (HR), mangster@unizd.hr

Francesca Di Garbo – Department of Languages, University of
Helsinki, Helsinki, Finland, francesca.digarbo@helsinki.fi

Eri Kashima – Department of Languages, University of Helsinki,
Helsinki, Finland, eri.kashima@helsinki.fi

Ricardo Napoleão De Souza – Department of Languages, University
of Helsinki, Helsinki, Finland, ricardofns@gmail.com

Kaius Sinnemäki – Department of Languages, University of Helsinki,
Helsinki, Finland, kaius.sinnemaki@helsinki.fi

Fabio Gasparini – Seminar für Semitistik und Arabistik, Freie
Universität Berlin, Berlin, Germany, fabiogsparini@gmail.com

Víctor Lara Bermejo – Dipartimento di Filologia, Università di Cadice, Cadice, victor.lara@uca.es

Antonietta Marra – Dipartimento di Lettere, lingue e beni culturali, Università di Cagliari, amarra@unica.it

Emanuele Miola – Dipartimento di Filologia Classica e Italianistica, Università di Bologna, Bologna, emanuele.miola@unibo.it

Adriano Murelli – Dipartimento di Studi Umanistici, Università di Torino, Torino, adriano.murelli@unito.it

Questo volume raccoglie le relazioni presentate al workshop Tipologia e sociolinguistica: verso un approccio integrato allo studio della variazione della Società di Linguistica Italiana, svoltosi on-line il 20 settembre 2020. Il tema del workshop è stato lo studio della variazione linguistica e interlinguistica, con l'obiettivo di gettare le basi per una nuova e più sistematica riflessione sulla possibile intersezione tra sociolinguistica e tipologia. Tutti i contributi sono stati sottoposti a doppia revisione anonima.

SILVIA BALLARÈ è assegnista di ricerca all'Università di Bologna. Si occupa di sociolinguistica dell'italiano, del rapporto tra variazione intra- e interlinguistica e di contatto linguistico.

GUGLIELMO INGLESE è FWO Junior Postdoctoral Fellow presso KU Leuven. I suoi interessi di ricerca si concentrano sui temi della linguistica storica, con particolare attenzione alle lingue indoeuropee antiche, e sulla linguistica tipologica.