

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 'rarità' 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 popolate 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 pronomi relativi (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.