

## Quante sono le lingue del mondo?

Il Signore disse: «Ecco, essi sono un solo popolo e hanno tutti una lingua sola; questo è l'inizio della loro opera e ora quanto avranno in progetto di fare non sarà loro possibile. Scendiamo dunque e confondiamo la loro lingua, perché non comprendano più l'uno la lingua dell'altro». Il Signore li dispersé di là su tutta la terra ed essi cessarono di costruire la città. Per questo la si chiamò Babele, perché là il Signore confuse la lingua di tutta la terra e di là il Signore li dispersé su tutta la terra.

Genesi 11, 1-9

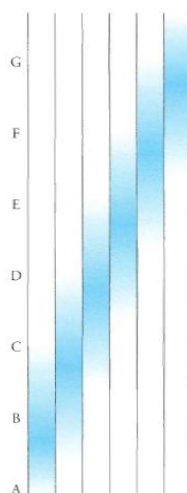
### 1. Quante sono le lingue del mondo?

Per quanto possa sembrare sorprendente, è molto difficile, se non impossibile, rispondere a questa domanda. Nella letteratura scientifica, riecheggiata con ignara sicumera dalle opere divulgative e dalla pubblicitaria, la stima ondeggia tra le 6.000 e le 12.000 lingue. Le stime più gettonate, comunque, sono tra le 6.000 e le 9.000, sebbene la moda attuale verta intorno alle 6.000. Qual è la ragione dell'imbarazzante divario? Confidando nel fatto che qualsiasi linguista sappia fare una somma, si può pensare che esistano zone della Terra ancora poco conosciute linguisticamente. Non è così: tutte le aree della Terra sono sufficientemente conosciute dal punto di vista linguistico.

Del resto, se ci ponessimo la stessa domanda riguardo alla ben familiare Italia, ci troveremmo davanti allo stesso imbarazzo: esistono tante classificazioni dei dialetti italiani quanti sono gli studiosi che si sono cimentati originalmente in questo compito; quella che segue è una delle tante proposte (la cartina riporta anche le varietà straniere):

In effetti, ci sono un paio di ragioni di natura contingente che sono parte del problema. La prima è che, dato l'elevato tasso di scomparsa delle lingue del mondo tra la fine del secolo scorso e l'inizio dell'attuale, è possibile che vengano censite lingue che non sono più parlate (o l'inverso: lingue date per morte, ma parlate ancora da uno sparuto numero di individui). Un'altra ragione, che incide di più nel conteggio, è che per la stessa lingua esistono **nomi varianti**: è facile contare due volte la stessa lingua sulla base di due denominazioni diverse attestate nella letteratura (si tenga presente che i linguisti – o gli antropologi – che si dedicano a questo tipo di ricerca contano per lo più su dati di seconda mano), oppure prendere per la stessa lingua due lingue diverse, ma affini, sulla base del fatto che condividono la stessa denominazione in qualche descrizione linguistica areale.

Tuttavia, nessuna delle ragioni appena discusse rendono conto di un divario che può ammontare, nei casi peggiori, al 100%. Di che cosa si tratta, allora? Il problema vero è intrinseco al concetto di *lingua*, che è estremamente elusivo. Il punto è che "in natura" non esiste nulla che corrisponda al nostro uso della parola 'lingua', i.e. un'individualità linguistica immediatamente riconoscibile. Quello che si trova sul territorio, come le ricerche dialettologiche dei due secoli scorsi hanno mostrato, è sì diversità linguistica, ma tale che non si lascia tagliare a fette, in quanto è di natura – normalmente – **graduale**. La figura seguente esemplifica il problema: le lettere sull'asse verticale stanno per comunità linguistiche diverse, gradualmente distanti a seconda dell'ordine alfabetico della lettera; le aree colorate stanno per i fatti linguistici comuni, quelle in bianco per i fatti linguistici diversi (come si vede, la comunità A condivide quasi tutti i fatti linguistici con la comunità B, un po' meno con la C, e via scemando):

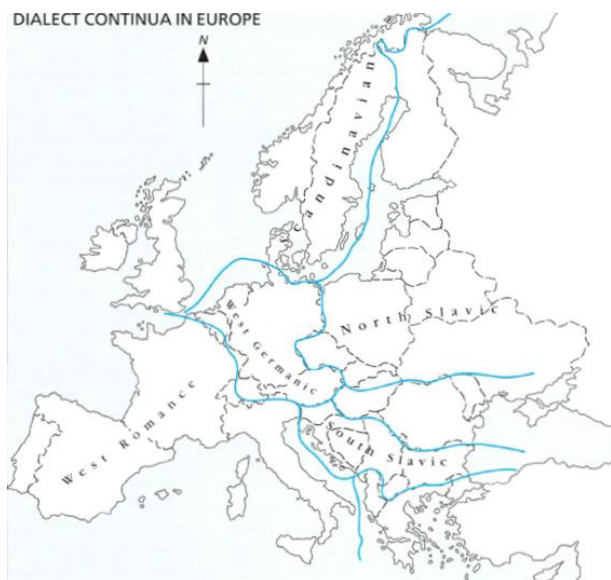


Questo fa sì che non sia affatto evidente dove finisca una lingua e ne cominci un'altra. Parlare di *dialetti*, piuttosto che di lingue, non migliora la situazione: quanto diverse debbono essere due varietà di lingua per essere considerate lingue diverse, invece che dialetti diversi della stessa lingua? In realtà, ci si potrebbe parimenti chiedere quanto due varietà linguistiche debbano essere diverse per essere considerate due dialetti diversi, invece che varianti dello



stesso dialetto. In una certa misura, il problema ammonta a chiedersi in quanti gruppi si possa dividere una sequenza numerica: per coppie, per terne, per cinquine, decine, ecc.? Evidentemente, ogni risposta è arbitraria.

La geografia linguistica, una branca della dialettologia, ha evidenziato l'esistenza dei *continua linguistici* (o *dialettali*), di cui viene data di seguito una mappa dell'Europa occidentale:

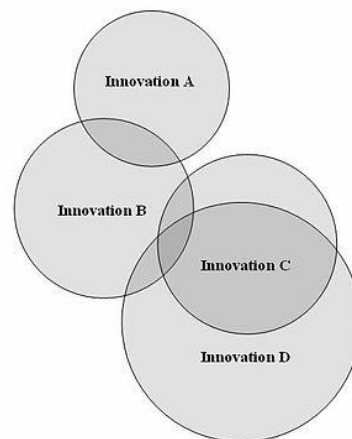


La realtà è ancora più sfumata di quanto la cartina, di fatto sommaria, indichi. I dialettologi hanno stilato delle mappe che rappresentano la distribuzione di singoli fatti linguistici, che hanno evidenziato il fatto che questi non sono discreti; anzi, sono non di rado distribuiti attraverso lingue diverse e persino famiglie linguistiche diverse, come la seguente cartina, relativa alla distribuzione delle *vocali anteriori arrotondate*, che ricorrono in *dialetti/lingue* diverse (piemontese, lombardo, francese, tedesco, olandese, frisone, danese, svedese, norvegese, finlandese), *sottofamiglie* diverse (Romanzo centrale e occidentale, Germanico occidentale e nordico) e *famiglie* diverse (Indoeuropeo, Ugrofinnico), mostra:



Come si spiega questo fenomeno? La risposta sta nella modalità in cui si attua il mutamento linguistico. Un mo-

dello che lo rappresenta è la 'teoria delle onde' (*Wellentheorie*) di J. Schmidt (in realtà, si tratta molto più modestamente di una metafora che coglie gli aspetti più basilari del problema). Il mutamento linguistico viene paragonato all'effetto di un sasso che cade al centro di uno stagno: genererà nell'acqua dei cerchi che diventeranno progressivamente più grandi e deboli. Questo allude al fatto che le innovazioni linguistiche che sorgono (per motivi ancora ignoti) si diffondono nel territorio (in realtà, nelle società) con maggiore e minore successo (per ragioni solo in parte comprese) a partire da un epicentro, come nella figura a lato.



A seconda delle innovazioni che una certa comunità linguistica accoglierà, si formerà la sua identità linguistica. Nella figura a lato si vede il tipico andamento della diversità linguistica nel territorio: immaginando, per amore di discussione, una condizione di partenza uniforme, quanti dialetti sono generati dalle innovazioni rappresentate? (La risposta è 9: 4 aree sono interessate da una innovazione; 4 da 2 innovazioni; infine, una è interessata da 3 innovazioni).

Questo semplice modello spiega perché, normalmente, la somiglianza/diversità linguistica è una funzione della distanza geografica. Si ricordi, però, che non sono i territori che parlano le lingue, ma le persone, o, meglio, i gruppi umani: la distanza è certo un ostacolo alla comunicazione, ma non l'unico. Esistono barriere di ogni tipo: fisico (catene montuose, mari, deserti, ecc.), sociale (segregazione in classi sociali diverse; relazioni conflittuali tra gruppi, ecc.), culturali (stili di vita molto diversi, ecc.), come anche rapporti privilegiati (si pensi, p.e., alla pratica dell'esogamia, che favorisce il contatto con altri gruppi specifici). Pertanto, non ci si deve aspettare dalla 'teoria delle onde' più di quanto non possa effettivamente dare, che è semplicemente questo: il mutamento linguistico, *a parità di condizioni* (i.e. in un mondo idealmente uniforme), ha effetto sulle comunità linguistiche sparse sul territorio in modo proporzionale alla loro distanza rispetto all'epicentro dell'innovazione.

Il problema, però, è che le condizioni non sono mai pari: per rimanere nella metafora, il sasso può essere un granello che increspa appena la superficie dell'acqua o un

macigno che provoca una sorta di onda anomala. La differenza dipende dai rapporti di potere che esistono tra epicentro (una certa comunità linguistica o, meglio, un gruppo di potere all'interno di essa) e periferia, in cui 'potere' va inteso in senso lato, includendo concetti come prestigio, ecc. Questo determina un'asimmetria nel mutamento linguistico che può essere molto marcata: alcune varietà linguistiche (comunità linguistiche) accolgono delle innovazioni da cui restano immuni altre varietà, le quali, piuttosto, *esportano* innovazioni. Quindi, esistono sassi e sassi, e lo stagno può essere molto asimmetrico.

Tutto ciò è mostrato dalla cartina dialettale dell'Italia, discussa in avvio: i dialetti individuati (si ignorino per ora le lingue straniere) sono, in senso stretto, delle pure astrazioni, in quanto le aree delimitate tramite colori diversi non sono affatto linguisticamente uniformi (invitiamo il lettore a verificarlo di persona nella sua area di competenza personale); d'altra parte, esse hanno una loro ragione di essere in un senso più astratto, che capita essere quello che corrisponde al modo in cui un parlante tipicamente categorizza la diversità linguistica, i.e. tramite prototipi: una varietà linguistica sarà giudicata tanto più toscana, per fare un esempio, quanto più condividerà i tratti tipici del fiorentino, il prototipo del dialetto toscano. La ragione dell'esistenza di questo prototipo è che è stata quella fiorentina (o, meglio, un gruppo sociale di essa) la comunità linguistica che ha più influenzato le comunità limitrofe, in cui 'limitrofo' è definito dallo spazio socio-politico condiviso: non è un caso che i confini dialettali riproducono, grosso modo, i confini politici storici, da cui derivano le regioni (e sottoregioni) italiane. Questi confini, inoltre, non sono impermeabili, come mostra il fatto (non notato nella cartina, che rappresenta una semplificazione della realtà) che un dialetto regionale sfuma nell'altro (i dialetti laziali nordoccidentali possono suonare quasi toscani a un orecchio non toscano, perché ne condividono alcuni tratti, come la gorgia, i.e. l'aspirazione delle occlusive sorde).

In realtà, i fatti sono ancora più complessi: come le ricerche sociolinguistiche del secolo scorso hanno mostrato, non ci sono due parlanti che parlano perfettamente allo stesso modo; anzi, esiste anche una variazione intraindividuale: nessuno parla allo stesso modo in tutte le situazioni.

D'altra parte, esistono, sebbene meno comunemente, le 'isole linguistiche', discontinuità linguistiche presenti nel territorio, i.e. comunità linguistiche in cui si parla una varietà linguistica che interrompe la continuità linguistica propria di un'area. In Italia, ad esempio, esistono isole linguistiche *albanesi* (Molise, Campania, Puglia, Sicilia, Calabria), *croata* (Molise: *slavisano* = *slavo molisano*), *greche* (Calabria, Sicilia, Puglia = *grico*), *catalana* (Alghero), *cimbria* (Luserna, Altopiano d. Sette Comuni, Giazza, Cansiglio; Bavarese? Nordico? "Gotico"?), *mochena* (valle d. Mocheni in Trentino; dialetto bavarese), *occitana* (Guardia Piemontese in Calabria), *siciliana* (Cilento meridionale, Campania), *ligure* (isola di S. Pietro in Sardegna: *tabarchino*).

*Come si formano le isole linguistiche?* Il processo di gran lunga più comune è una discontinuità etnica: i grandi movimenti di popoli (o anche piccoli, come nel caso delle comunità zingare), specie in fase iniziale, provocano discontinuità linguistiche. L'avanzata delle popolazioni di lingua indoeuropea in Europa ha emarginato le comunità linguistiche precedenti, come quella basca e quelle caucasiche, alle estremità dell'Europa, creando una discontinuità in quello che doveva essere un continuum linguistico. L'isola linguistica basca è ciò che rimane in Europa di uno stato linguistico pre-neolitico. A sua volta, l'isola ungherese, in mezzo a un continuum linguistico essenzialmente slavo, è il frutto di emigrazioni posteriori a quella indoeuropea.

Esistono anche le 'penisole linguistiche', discontinuità linguistiche presenti nel territorio che sono continue rispetto a varietà presenti in territorio straniero; si tratta quindi di varietà linguistiche che "trasbordano" attraverso i confini nazionali, a riprova del fatto che nessuna frontiera è impermeabile agli influssi linguistici e che i confini politici non coincidono necessariamente con quelli linguistici e culturali. In Italia esistono penisole linguistiche *franco-provenzali* (Valle d'Aosta e valli confinanti del Piemonte), *occitana* (provenzale) (Alpi occidentali), *tedesca* (Monte Rosa, Val Formazza, Alto Adige), *slovena* (Provincia di Trieste, alcuni comuni della provincia di Gorizia, Slavia veneta, Val Canale).

Stando così le cose, come possiamo rispondere allora alla domanda iniziale? A rigore, la risposta dovrebbe essere la seguente: *ci sono tante lingue/dialetti quante persone al mondo*. Si noti che non si tratta di una boutade provocatoria, ma di un'asserzione corretta, da un certo punto di vista. D'altra parte, oscura, se presa alla lettera, un'ovvia verità: che queste circa sette miliardi di varietà linguistiche non sono ugualmente diversificate tra loro, potendosi trattare, a seconda dei casi, di varietà che condividono il 99,9 % dei fatti linguistici (ammesso che fossimo in grado di fare valutazioni del genere) e varietà che condividono solo fatti linguistici derivati dalla dotazione linguistica innata della nostra specie. In termini pratici, questo significa che quando si parla della grammatica e dizionario di una "lingua", queste generalizzazioni saranno tanto più concrete quanto minore è la comunità linguistica considerata: una descrizione della pronuncia dell'italiano parlato a Firenze è più concreta della descrizione dell'italiano parlato in Toscana, che è a sua volta più concreta della descrizione della pronuncia dell'italiano parlato in Italia.

La conclusione di questo lungo discorso è che, *per quanto possa suonare paradossale a un non addetto ai lavori, le 'lingue', nell'uso comune del termine, non esistono*; esiste la *diversità linguistica*, nello stesso tempo continua e discontinua. La concezione di lingua di senso comune deriva dalle lingue nazionali, che sono entità semiartificiali: si tratta di una varietà di lingua appartenente a una comunità linguistica specifica che è stata estesa arbitrariamente, in un'estensione di tempo variabile, a un intero territorio, in genere nazionale, sovrapponendosi, in qualche modo, alla varietà locale. Questo è, ad esempio,

quello che è avvenuto in Italia in tempi relativamente recenti, non più di un secolo e mezzo fa, come mostra la presenza nel nostro Paese di dialetti ancora forti. Per questo, all'estremità opposta del criterio di giudizio riguardo al numero delle lingue presenti in una certa area, c'è quello che conterebbe la presenza di una sola lingua in Italia, l'"italiano". Questa è, con tutta probabilità, la risposta che un italiano medio darebbe al problema. I romanisti si regolano in modo leggermente diverso: isolate le lingue straniere (i.e. quelle che sono parlate anche al di fuori dei confini italiani, come ladino, tedesco, sloveno, ecc.), considerano tutte le varietà linguistiche presenti in Italia come 'dialetti italiani', ad eccezione del sardo e del friulano, giudicati troppo divergenti dall'italiano comune per essere considerati delle sue varietà. Ovviamente, questa scelta è fondata su fatti oggettivi: il friulano e ancor più il sardo sono le varietà linguistiche romanze presenti in Italia più divergenti dalle altre; tuttavia, non è immune da un certo grado di convenzionalità: a quale grado di divergenza scatta la differenza tra dialetto di una stessa lingua e lingua diversa? Chiaramente, ogni risposta è arbitraria.

Tuttavia, ci si potrebbe chiedere come si regolino normalmente i linguisti che elaborano i dati geolinguistici. Ebbene, la risposta è per un verso molto semplice, per altri molto complessa. È *semplice* in quanto si opera per riduzione: le varietà considerate strettamente imparentate vengono considerate una sola lingua ('dialetti della lingua L'). Questa pratica, però, come il lettore attento avrà certamente notato, incorre nel solito problema: come si fa a stabilire in modo univoco quando due varietà sono 'strettamente imparentate'? Il problema è che la somiglianza è graduale, non discreta. La risposta è che ogni ricercatore si regola un po' a naso. Una conseguenza è che le descrizioni areali non sono necessariamente uniformi: a seconda di chi le ha elaborate, la generosità nel riconoscere individualità a singole varietà può variare in un certo grado. Pertanto, le stime mondiali sul numero delle lingue non si basano su analisi uniformi. Quanto questo incida nella precisione del conto è difficile da dire.

C'è un altro aspetto del problema, che può essere ancora più rilevante. Il concetto di *dialetto* è in effetti ambiguo: da un lato indica una varietà linguistica strettamente imparentata con un'altra ('A e B sono dialetti della stessa lingua'); dall'altro indica inferiorità sociale ('A è un dialetto, non una lingua'). Il punto è che questa ambiguità si può riflettere anche negli studi linguistici areali: il senso di dialetto inteso come varietà subordinata a una lingua è tipico di molti stati unitari di lunga tradizione; non lo è invece di quelli di recente unificazione o di società preindustriali (se non preneolitiche), dove tutte le varietà sono alla pari. Questo fa sì che l'assegnazione di status di 'lingua', piuttosto che di 'dialetto', a una varietà linguistica sia effettuata in modo più generoso nelle aree in cui non esiste una tradizione unitaria rispetto a quelle fortemente centralizzate dal punto di vista politico.

Per cogliere questa differenza è di grande utilità il concetto di *lingua tetto* (elaborato da H. Kloss, un sociolin-

guista tedesco): 'Una lingua è l'insieme delle varietà linguistiche costituito da una varietà standard più tutte le varietà da essa coperta'. Il concetto di lingua tetto/copertura allude al fatto che nelle società di tipo lingua *cum dialectis* i dialetti sono subordinati alla lingua standard e ne subiscono l'influenza (che porta, eventualmente, alla loro riduzione o, nei casi estremi, alla loro scomparsa). Applicando questo concetto alle diverse comunità linguistiche, si ottiene una dicotomia interessante:

+ LINGUA TETTO	-
dialetti italiani, inglesi, francesi, ecc.	dialetti zingari, italoalbanesi, ecc.

La tabella mostra che, diversamente dai dialetti italiani, inglesi e francesi (per limitarci ad alcuni esempi), i dialetti zingari non hanno una lingua tetto, i.e. una varietà linguistica considerata superiore alle altre (per la ragione che la loro identità comune si basa su elementi culturali, non politici). Lo stesso vale per i dialetti albanesi presenti nel nostro territorio, ma per ragioni diverse: esiste certamente una lingua standard albanese, ma le comunità linguistiche albanesi in Italia ne sono indipendenti, i.e. non la utilizzano come lingua alta. La ragione è che per gli usi "alti" (a Scuola e all'Università, negli usi amministrativi, ecc.) utilizzano l'italiano. Si tratta, quindi, di comunità diglossiche bilingui.

Sempre a Kloss si deve una proposta che esplicita, seppure in modo relativistico, i criteri in base ai quali si considera una varietà linguistica come *lingua* ovvero come *dialetto*. I parametri che utilizza sono l'*elaborazione* e la *distanziamento* di una lingua. Il grado di elaborazione di una lingua è definita in base a due parametri, gli *argomenti* che è in grado di veicolare e il *livello* massimo a cui lo può fare ('livelli di sviluppo'):

	ARGOMENTI	LIVELLI SVILUPPO
-	storia e tradizione locale	scuola elementare
	cultura generale	scuola secondaria
+	scientifici e tecnologici	università

Questo significa che una lingua è massimamente elaborata se può veicolare anche contenuti di natura tecnico-scientifica a livello universitario; una lingua è invece minimamente elaborata se è grado di veicolare solo contenuti legati alla storia e tradizioni locali e a non più del livello di scuola elementare. Ne consegue che l'italiano è una lingua massimamente elaborata, mentre molte lingue etniche lo sono minimamente: alcune vengono utilizzate per l'insegnamento elementare, ma non oltre; altre fino alla Scuola secondaria, ma non nell'insegnamento universitario (oppure solo nelle Facoltà legate alle tradizioni religiose o culturali del posto). I dialetti nelle comunità di tipo *lingua cum dialectis* hanno un grado di elaborazione ancora inferiore. L'inglese è oggi la lingua che al momento ha il più alto grado di elaborazione, fornendo la nomenclatura tecnico-scientifica a tutte le altre lingue del mondo, come lo erano state in Occidente (con le dovute differenze areali e temporali) il greco e il latino. Si noterà

che il parametro del livello di sviluppo, così come è formulato, non è universalmente applicabile, ma sembra confezionato su misura per le società occidentali (o occidenteggianti) contemporanee; tuttavia, considerato il fatto che la sua finalità non è antropologica e che il modello occidentale è dominante nel mondo contemporaneo, essa ritiene la sua utilità tassonomica.

Il grado di distanziamento di una varietà linguistica, ovvero quanto è diversa da un'altra varietà, è l'altro parametro utilizzato comunemente, da addetti ai lavori come dai profani, per decidere se due varietà sono dialetti di una stessa lingua o due lingue diverse. Il problema, come si ricorderà, è che non c'è un modo univoco per stabilire l'affinità tra varietà linguistiche (anche se si potesse effettuare un conteggio delle differenze, come si potrebbe stabilire il peso di ciascuna differenza?) né, quello che più conta, è possibile individuare un livello soglia non arbitrario. I criteri che sono stati utilizzati dai diversi tipi di linguisti che si scontrano con questi problemi (storici, dialettologi, etnolinguisti, sociolinguisti) sono diversi, a volte utilizzati in concomitanza, a volte in alternativa.

Il primo è la *parentela genealogica*, definita in base ai criteri sviluppati nell'ambito della grammatica storico-comparativa. È in base a questo criterio, ad esempio, che i romanisti assegnano al sardo, al friulano e al ladino uno stato diverso da quello degli altri dialetti romanzi presenti in Italia.

Un altro criterio è la *differenza strutturale*, i.e. si decide solo in base alle caratteristiche esistenti, senza tenere conto della derivazione storica. Entrambi questi criteri, però, si scontrano col problema della gradualità, già visto ripetutamente in questa trattazione: parentela e somiglianza strutturali non sono discreti.

Un altro criterio ancora utilizzato è la *coscienza dei parlanti*: invece di addentrarsi in analisi faticose e poco conclusive si possono interrogare gli informanti locali per stabilire i rapporti tra varietà vicine. Si tratta di una procedura di tipo psicologico, che ha il merito di essere olistica. Purtroppo, però, la coscienza dei parlanti, per quanto utile, non è affidabile, perché può basarsi su ogni tipo di pregiudizio o ideologia. Ad esempio, in Cina si ha un atteggiamento molto radicato a ritenere che vi si parli una sola lingua, sebbene divisa in diversi dialetti; in realtà, alcuni di questi supposti dialetti, come il cantonese, sono largamente incomprensibili per chi conosce solo il cinese standard, il pechinese (divergendo tra loro più di quanto non facciano l'italiano e lo spagnolo, per fare un esempio).

Infine, un altro criterio largamente utilizzato è, appunto, *l'intercomprensibilità*: se due varietà sono mutuamente comprensibili sono dialetti della stessa lingua, altrimenti sono lingue a se stanti. Tuttavia anche questo criterio risulta problematico, per almeno due ragioni. La prima è che anche l'intercomprensibilità non è una faccenda tutto/nulla ma graduale. I risultati possono essere paradossali; infatti, mentre i parlanti dell'*italiano* e dello *spagnolo*, considerate lingue diverse, possono comprendere una porzione considerevole dell'altra lingua, i parlanti del

*lombardo* e del *siciliano*, considerati dialetti dell'italiano, possono incontrare barriere grandissime alla mutua comprensione. Un altro problema è la soggettività dei giudizi: persone diverse danno giudizi diversi riguardo alla comprensione di parlanti di varietà affini. Un particolare caso a questo riguardo sono le asimmetrie di giudizio, di natura questa volta collettive: che cosa dire del caso in cui i parlanti di una comunità A asseriscono di capire abbastanza bene i parlanti della comunità B, ma non l'inverso? (Un esempio di questo fenomeno è rappresentato in Europa dai portoghesi, dalla parte dei comprendenti, e dagli spagnoli, dalla parte dei non comprendenti; si noti che lo stesso fenomeno non sembra verificarsi in America tra brasiliani e ispanofoni).

Comunque sia, mettendo da parte tutte le riserve connesse alla determinazione della distanziamento tra varietà di lingua diverse, la proposta di Kloss è che una varietà di lingua sia tanto più suscettibile di essere considerata una lingua a sé quanto più è elaborata o quanto più è distante dalle altre varietà. La tabella seguente illustra il sistema tramite alcuni esempi che riguardano le principali varietà parlate in Italia:

		ELABORAZIONE		
		+	±	-
DISTANZIAMENTO	+	tedesco "francese" sloveno croato greco catalano albanese		varietà zingare dial. italoalbanesi
	±		sardo friulano ladino	
	-			dial. italiani

Ne risulta che le varietà tedesche, "francesi" (provenzali o franco-provenzali), slovene, croate, greche, catalane e albanesi parlate in Italia sono considerate lingue a se stanti, invece che dialetti dell'italiano, per il loro alto grado di distanziamento da qualunque varietà di italiano; hanno inoltre un alto grado di elaborazione perché sono lingue nazionali in altri Paesi. All'estremo opposto si situano i dialetti italiani, che non sono lingue né per grado di elaborazione (molto basso) né per distanziamento (non considerata abbastanza accentuata in nessun caso). In mezzo si situano varietà italiane come il sardo e il friulano, o una varietà come il ladino, non esclusivamente parlato in Italia, che sono intermedie sia come grado di elaborazione (esistono dizionari e grammatiche di riferimento, un'ortografia specifica, ecc.) sia come distanziamento (sono distanti dall'italiano medio, ma non tanto quanto le lingue prima discusse, tutte appartenenti a un'altra famiglia linguistica eccetto il catalano e il "francese", comunque più distanti dall'italiano del sardo o del friulano o del ladino).

I due parametri possono essere in conflitto: quando questo accade è in genere il grado di elaborazione che decide riguardo all'autonomia di una varietà rispetto a un'altra, come il caso del tedesco, dei dialetti tedeschi e del neer-

landese (parlati nell'area Germania/Olanda/Belgio fiammingo) evidenza (E = elaborazione; D = distanziamento):

		E		
		+	±	-
D	+			
	±			
	-	neerlandese		dialetti tedeschi

Il neerlandese è considerato una lingua a sé, piuttosto che un dialetto del tedesco, non per *distanziamento* (i dialetti neerlandesi parlati in prossimità del confine olandese-tedesco risultano largamente comprensibili ai tedeschi prossimi al confine, i.e. si tratta di un continuum dialettale), ma per *elaborazione*: è lingua nazionale in Olanda e una delle lingue nazionali del Belgio. È evidente che l'autonomia politica favorisce l'elaborazione di una lingua. Al contrario, i dialetti tedeschi non hanno autonomia linguistica in quanto hanno un basso grado di elaborazione e un basso grado di distanziamento dal tedesco standard.

Per concludere questa lunga discussione, quante sono allora le lingue del mondo, "in pratica"? La risposta è che *dipende dai criteri utilizzati*: più larghe sono le maglie della retina del setaccio utilizzato, più lingue vengono individuate e conteggiate; più strette e meno lingue saranno conteggiate. Il problema delle stime linguistiche disponibili è che esse si basano su dati spuri: non c'è nessuna garanzia, anzi, è fortemente improbabile, che ciascun ricercatore che ha lavorato su una singola area abbia adottato gli stessi criteri degli altri. La caratterizzazione della diversità linguistica presente in Italia, un'area non esattamente marginale, ne è una dimostrazione lampante. Questo spiega il divario, a volte impressionante, nelle stime delle lingue del mondo. È importante sottolineare, infine, che non si vuole qui intendere che queste stime siano inutili sciocchezze; al contrario, hanno un'utilità in quanto indicative dell'ordine di quantità del fenomeno linguistico. L'importante è non prenderle troppo sul serio (specialmente se formulate approssimate all'unità).

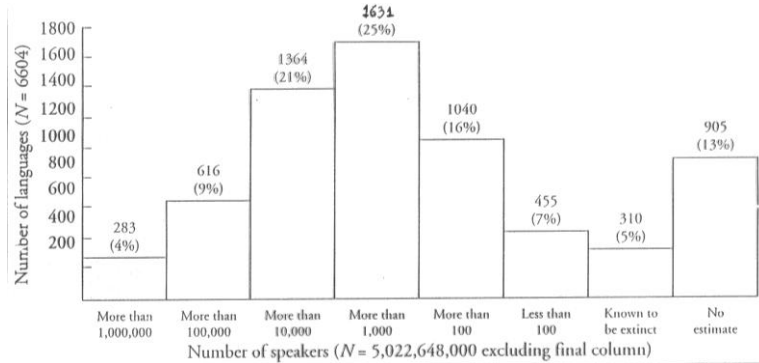
## 2. Da quante persone sono parlate le lingue del mondo?

Le prime 15 lingue per numero di parlanti madrelingui sono le seguenti (l'italiano in molte stime viene immediatamente dopo il telugu, con i suoi circa 65.000.000 di parlanti, che ne fanno comunque una delle lingue più parlate al mondo):

		POPOLAZIONE	%
1	cinese mandarino	885.000.000	15
2	inglese	322.000.000	5,4
3	spagnolo	266.000.000	4,5
4	bengali	189.000.000	3,2
5	hindi	182.000.000	3
6	portoghese	170.000.000	2,8
7	russo	170.000.000	2,8
8	giapponese	125.000.000	2,1
9	tedesco	98.000.000	1,6
10	cinese wu	77.175.000	1,3
11	giavanese	75.500.800	1,2
12	coreano	75.000.000	1,2
13	francese	72.000.000	1,1

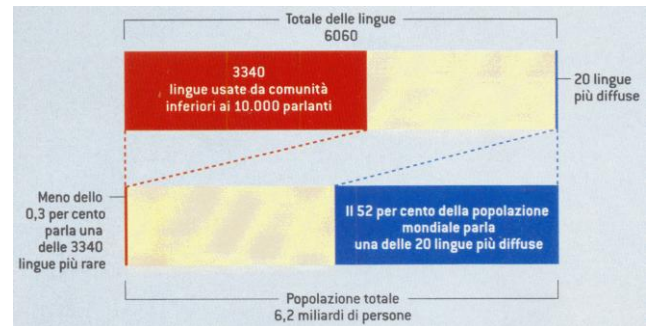
14	vietnamita	66.897.000	1,1
15	telegu	66.350.000	1,1

Queste cifre, però, non ingannino: l'esistenza di lingue parlate da decine di milioni di persone è l'eccezione, non la regola. Il grafico seguente mostra la distribuzione del numero dei parlanti per lingua:



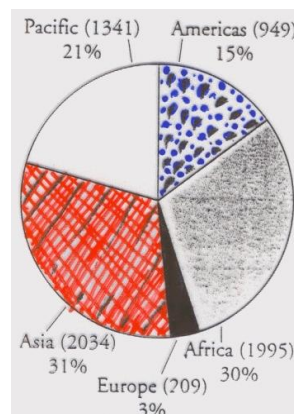
Come si vede, solo il 4% delle lingue è parlata da più di un milione di persone; la gran parte delle lingue, l'87%, è parlata da meno di 100.000 persone (un 13% rappresenta lingue ormai estinte o di cui non si hanno stime affidabili, segno in genere che sono a rischio di scomparsa).

Il numero dei parlanti per lingua può essere enormemente asimmetrico, come mostra il grafico seguente, in cui le 20 lingue più diffuse sono parlate dal 52% della popolazione mondiale, mentre ci sono 3.340 lingue (poco più della metà) parlate da meno di 10.000 parlanti (lo 0,3% del totale dei parlanti):



## 3. Come sono distribuite le lingue nel mondo?

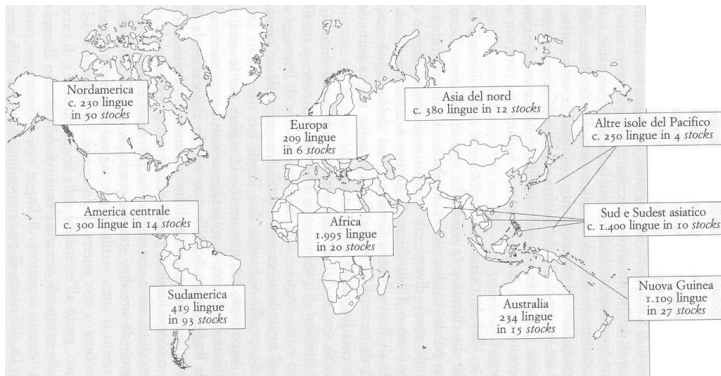
Anche la distribuzione delle lingue (o, per meglio dire, della *diversità linguistica*) per aree della Terra è assai asimmetrica, come il grafico a lato evidenzia.



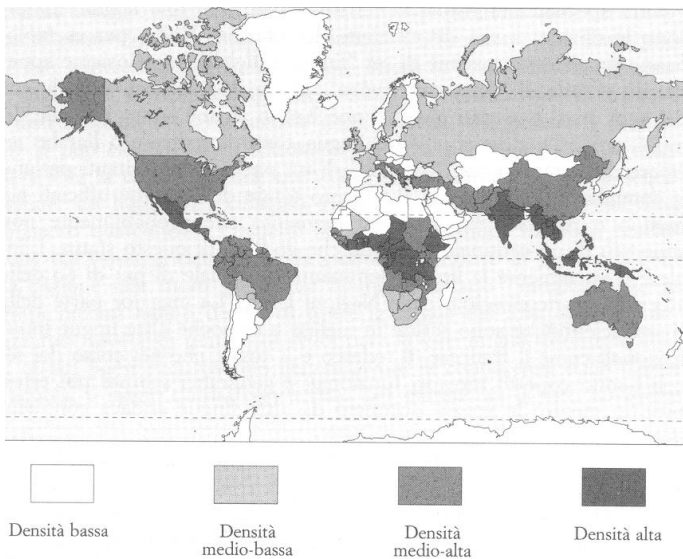
In assoluto, l'Africa risulta il continente con maggiore diversità linguistica (pari circa a quella dell'Asia) in rapporto al numero degli abitanti (con una popolazione attuale di c. un miliardo di abitanti – ma di 221 milioni fino al 1950 – circa un quarto rispetto a quella dell'Asia). I dati più interessanti, tuttavia, sono quelli estremi: l'Europa, da una parte, – un'area molto popolata ma con la più bassa diversità linguistica, in assoluto e relativamente al numero degli abitanti – e l'Oceania, dall'altra – con la più alta diversità linguistica in rapporto alla popolazione. Nella tabella seguente vengono riportate le percentuali di lingue, divise per continente, con un numero di parlanti inferiore al numero indicato:

Continente	< 150	< 1.000	< 10.000	< 100.000	< 1.000.000
Africa	1,7	7,5	32,6	72,5	94,2
Asia	5,5	21,4	52,8	81	93,8
Europa	1,9	9,9	30,2	46,9	71,6
Nordamerica	22,6	41,6	77,8	96,3	100
Centramerica	6,1	12,1	36,4	89,4	100
Sudamerica	27,8	51,8	76,5	89,1	94,1
Australia/Pacifico	22,9	60,4	92,8	99,5	100
Mondo	11,5	30,1	59,4	83,8	95,2

Una mappa più dettagliata in quanto ad aree geografiche è la seguente:

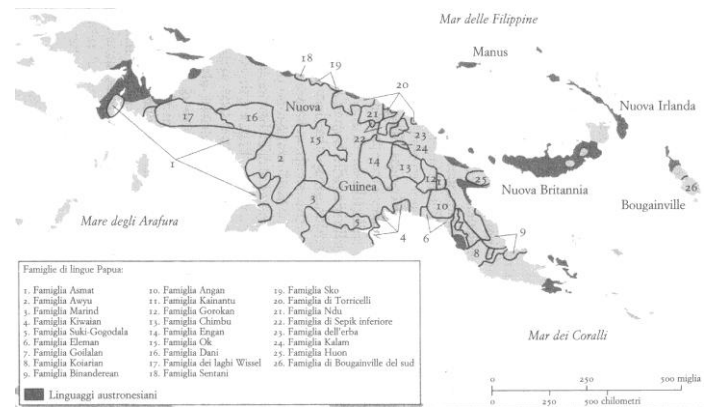


La seguente mostra invece la diversità linguistica relativa alle principali nazioni:



Come si spiegano queste forti asimmetrie? Proprio i casi dell'Europa, da una parte, e dell'Oceania, dall'altra, forniscono la chiave per rispondere a questa domanda. L'Eu-

ropa è il continente che più e di tutti ha conosciuto il fenomeno della formazione di grossi stati centralizzati, dapprima con l'Impero Romano, che ha uniformato linguisticamente l'Europa occidentale con il latino e l'Europa orientale col greco, poi con la formazione degli stati attuali, che hanno imposto una lingua comune. In altre aree non si è verificato un fenomeno dello stesso grado (a parte che in Asia, ma non con la stessa sistematicità). Perché l'Oceania è così differenziata linguisticamente nonostante la bassissima densità demografica (conta solo c. 36.000.000 di abitanti, ma va anche tenuto conto che circa la metà della popolazione vive in Australia)? La ragione è inversa a quella dell'Europa: a parte l'Australia e la Nuova Zelanda, di formazione comunque relativamente recente e in cui l'inglese si è imposto a spese delle lingue autoctone, non esistono in Oceania Stati di grande estensione e di lunga tradizione. La ragione è fondamentale geografica: da un lato, si tratta di un territorio disseminato in forma di isole su un vastissimo oceano (che costituisce una barriera all'uniformità socio-politica e quindi linguistica); dall'altro, la gran parte della diversità linguistica si trova concentrata in un'unica isola, la Nuova Guinea (divisa politicamente tra Indonesia e Nuova Guinea Papua), che può essere considerata a buon diritto il paradiso della diversità linguistica, con le sue c. 1.100 lingue diverse censite in rapporto a c. 7.500.000 di abitanti. Si noti che non si tratta di dialetti: sono più o meno tutte mutuamente incomprensibili e sono divise in 26 sottofamiglie (senza contare le lingue austronesiane, di provenienza asiatica):



Come si spiega l'anomalia papuana? Anche qui la ragione sta nella – sbalorditiva – frammentazione sociale che ha caratterizzato l'isola fino a tempi recenti, caratterizzata da catene montuose all'interno che hanno formato una barriera alla comunicazione dei popoli autoctoni, che, inoltre, abitano continuamente l'isola da almeno 40.000 anni (alcune stime, però, retrodatano l'arrivo dei papuani di almeno 10.000 anni). Si immagini l'effetto del mutamento linguistico non condiviso tra gruppuscoli di individui per un arco di tempo così rilevante.

In realtà, la situazione della Nuova Guinea Papua non dovrebbe stupire alla luce della storia della nostra specie; semmai, è la situazione dell'Europa ad essere anomala. La ragione è che i grandi gruppi sociali hanno cominciato ad affermarsi solo "recentemente" in seguito alla rivolu-

zione neolitica, che si è avuta, dapprima in pochissime aree del mondo, a cominciare da c. 10.000 anni fa, con l'addomesticamento e la selezione di alcune specie vegetali e animali. Perciò, c'è ragione di ritenere – su basi deduttive dettate dall'osservazione dei gruppi con un'economia ancora pre-neolitica – che per tutta la storia dell'umanità la situazione sociale e quindi linguistica sia stata molto più simile a quella della Nuova Guinea Papua che a quella dell'Europa attuale.

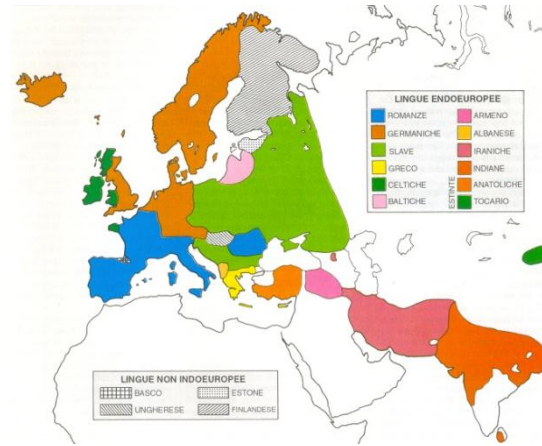
#### 4. Quali relazioni di somiglianza intercorrono tra le lingue?

Appurato che il linguaggio umano ha un'indubbia vocazione per diversificarsi, risultato a cui concorre il mutamento linguistico, questo non significa che le lingue del mondo siano equidistanti tra loro; al contrario, la scoperta che alcune lingue mostrano corrispondenze sistematiche tra forma e contenuto (i.e. parole o altre forme simili per pronuncia e per significato) con alcune lingue ma non con altre è stato il risultato più importante della linguistica del XIX secolo. Un esempio di questo tipo di relazione è illustrato dai dati seguenti, il paradigma del verbo 'portare' in alcune lingue antiche, alcune delle quali molto distanti geograficamente (come il sanscrito, antica lingua dell'India settentrionale) o per attestazione temporale (tra le prime attestazioni del sanscrito e quelle del gotico intercorrono almeno dai 15 ai 20 secoli):

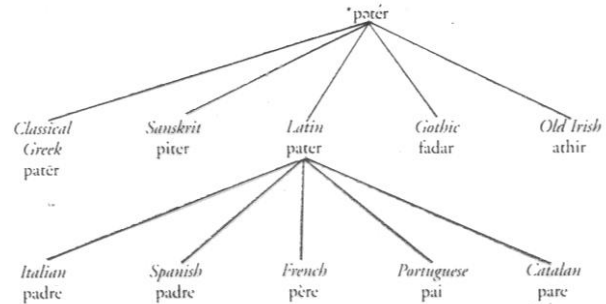
	SANSKRITO	GRECO CLASSICO	LATINO	IRLANDESE ANTICO	GOTICO
io porto	bhár-āmi	phér-ō	fer-ō	bir-u	bair-a
tu porti	bhár-asi	phér-eis	fer-s	bir-i	bair-is
egli porta	bhár-ati	phér-ei	fer-t	ber-id	bair-ith
noi portiamo	bhár-āmas	phér-omen	fer-imus	ber-mi	bair-am
voi portate	bhár-atha	phér-ete	fer-tis	ber-the	bair-ith
essi portano	bhár-anti	phér-ousi	fer-unt	ber-it	bair-and

Come si vede, le somiglianze non possono essere casuali (se si prendono due lingue qualsiasi del mondo, esse mostreranno un certo numero di somiglianze di forma e contenuto del tutto fortuite), in quanto (1) le corrispondenze di suono sono regolari (ad esempio, a /bh/ del sanscrito corrispondono /ph/ in greco, /f/ del latino e /b/ in gotico e irlandese, tutti fonemi labiali); (2) seguono visibilmente lo stesso modello morfologico (stesse categorie flessive e desinenze molto simili). La spiegazione che si è data, e continua a darsi, di questo tipo di affinità è la *parentela genealogica*, i.e. si tratta di un'eredità comune del passato: tutte queste lingue derivano da un antenato comune, la 'lingua madre'. In termini più reali, questo significa che quelle che nella fase storicamente osservabile si mostrano come comunità linguistiche diverse erano una volta (in cui 'una volta' va specificato) un sola comunità. La famiglia in questione è quella Indoeuropea (così denominata perché le lingue che vi appartengono sono sparse tra Europa e India (in realtà anche più a oriente, se si considera

una lingua ormai estinta, il tochario, parlato nella parte occidentale dell'odierna Cina):

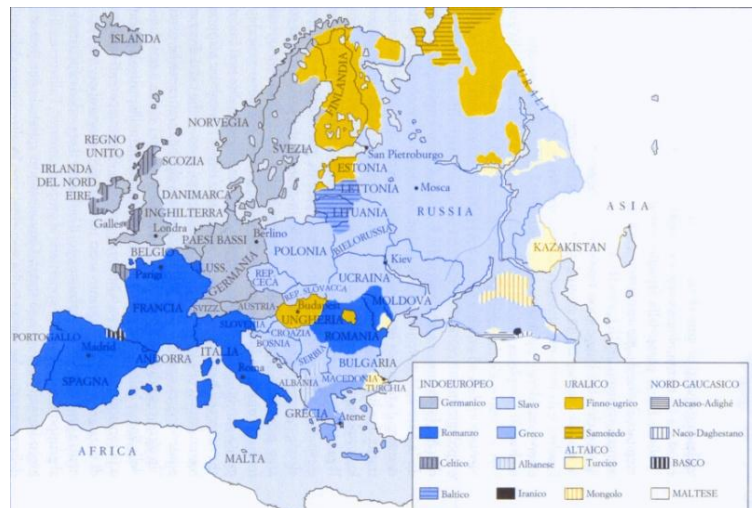


Questo ha permesso di stilare degli alberi genealogici (come quello qui sotto) che rappresentano le fasi della diaspora della comunità originaria protrattasi per molte migliaia di anni (in subordine, è mostrata la diaspora del latino, che ha prodotto le lingue romanze):



Un altro risultato, come mostra la stessa figura, è che è stato possibile, tramite tecniche comparative, ricostruire alcuni elementi di questa lingua originaria, non attestata (nella figura si tratta della parola per 'padre').

Se la gran parte delle lingue d'Europa appartiene alla famiglia Indoeuropea (specialmente a una delle sottofamiglie romanza, germanica e slava), ve ne sono altre di diversa origine: ugrofinnica (p.e., finlandese e ungherese), caucasica (come il georgiano), semitica (l'arabo maltese) e, molto marginalmente, il turco, lingua asiatica.





Una lingua non indoeuropea esclusivamente europea è invece il basco, che risale alla lingua dei più antichi abitanti conosciuti dell'Europa.

Esistono numerose altre famiglie linguistiche oltre a quella indoeuropea, che sarebbe troppo lungo e tedioso trattare in dettaglio (si rimanda per questo a una delle trattazioni specifiche). A colpo d'occhio, la parentela genealogica tra le lingue umane è rappresentata nella cartina seguente:



Si notano, da una parte, la relativa uniformità delle Americhe e dell'Europa; dall'altra, la grande difformità dell'Africa, in cui si trovano quattro macrofamiglie irriducibilmente autonome, almeno alla luce delle ricerche condotte fino ad oggi (non si considerano le lingue indoeuropee risultato delle colonizzazioni recenti e la lingua austronesiana parlata in Madagascar, risultato di colonizzazione secolare da parte di genti del Sud-est asiatico).

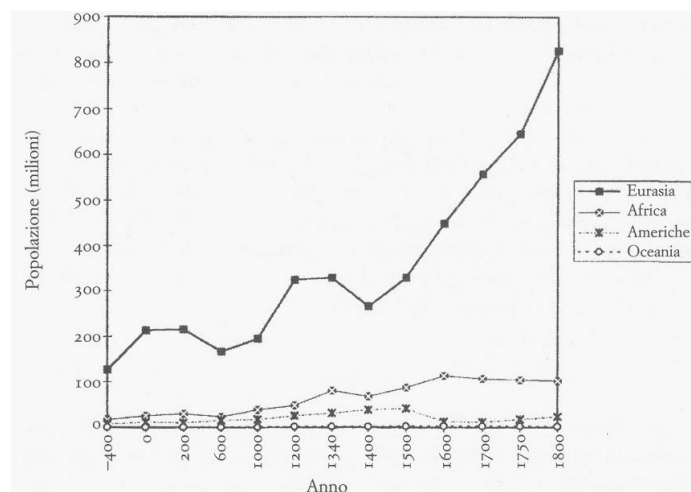
Non sorprende che anche la distribuzione dei parlanti per famiglie linguistiche sia assai diseguale. Le maggiori famiglie linguistiche per numero di parlanti sono le seguenti:

	LINGUE	PARLANTI
INDO-EUROPEO	386	2.500.000.000
SINO-TIBETANO	272	1.088.000.000
AUSTRONESIANO	1212	269.000.000
AFRO-ASIATICO	338	250.000.000
NIGER-CONGO	1354	206.000.000
DRAVIDICO	70	165.000.000
AUSTRALIANO	262	30.000

Il record della famiglia indoeuropea deriva, ovviamente, dal colonialismo di alcuni Stati europei: Gran Bretagna, Francia, Spagna, Portogallo, Russia. Il record negativo delle lingue aborigene australiane è l'altro lato della medaglia dello stesso fenomeno: alcune lingue europee si sono espanse a discapito di molte lingue aborigene. L'inglese si è affermato a spese delle lingue australiane, di quelle austronesiane della Nuova Zelanda, in parte delle lingue bantu e khoisan dell'Africa australe e delle lingue amerindiane del Nordamerica (Alaska, Stati Uniti e Canada). Lo spagnolo si è affermato sulle lingue amerindiane del Sud, Centro e parte del Nord America (Messico); il portoghese si è espanso anch'esso a spese delle lingue

amerindiane dell'odierno Brasile. Il francese ha soppiantato alcune lingue amerindiane della parte occidentale del Canada. Dopo l'Australia, è proprio l'America il continente che attesta maggiormente la scomparsa delle lingue autoctone. In Nord America, in particolare, con l'eccezione del navaho, che conta su di un mezzo milione di parlanti, tutte le lingue amerindiane sono parlate da un numero molto limitato di parlanti, in costante diminuzione. Nell'America latina le cose vanno un po' meglio, in particolare in quella del Sud, dove ci sono due lingue parlate da milioni di persone: il guaraní e, ancor più, il quechua, entrambe della sottofamiglia Andino-equatoriale. Va sottolineato, però, che la situazione del guaraní è migliore di quella del quechua, nonostante il minor numero di parlanti: il quechua è una lingua etnica, parlata solo da persone di quella etnia; viceversa, il guaraní è parlato in Paraguay anche da persone di origine europea. Pertanto, il suo futuro sembra più roseo.

Alcuni dati demografici, le stime della popolazione mondiale per continente tra il 400 a.C. e il 1800 d.C., evidenziano la natura di questo decadimento:



Come il grafico mostra, mentre in Eurasia c'è stata un'impennata demografica dal 1500, dopo la scoperta europea del Nuovo Continente è cominciata la decadenza demografica in America (l'Africa è cresciuta poco e l'Oceania è rimasta stabile).

### 5. Lingue che muoiono

È stato previsto da più parti che durante questo secolo scompariranno dal 50 al 60% delle lingue del mondo. Ma perché le lingue "muoiono"? Si ricordi che le lingue non esistono realmente, quindi non possono "morire" in senso letterale. Quello che esiste in termini reali sono i parlanti e le comunità linguistiche. Questo significa che intere comunità si estinguono? Non necessariamente (e, con tutta probabilità, nemmeno tipicamente). Certamente la decadenza demografica può essere una parte del problema (e questo deve essersi verificato spesso nel corso della storia della nostra specie), ma la ragione principale è semplicemente che una lingua cessa di essere parlata dal gruppo che la parlava. Per quale ragione? Una ragione frequente è la stessa per cui alcune innovazioni si diffon-

dono sul territorio a spese di altri elementi linguistici, come abbiamo visto: differenze di prestigio. Ad esempio, in America quasi tutte le lingue autoctone sono in crisi perché vengono percepite, anche dall'interno delle stesse comunità che le parlano, come lingue etniche, quindi inferiori. Una tipica trafila di morte di una lingua è la seguente: (1) il padre trasmette la lingua etnica al figlio, (2) il quale impara anche la lingua ufficiale di origine straniera (tipicamente a scuola) e assorbe gli atteggiamenti negativi nei confronti della propria lingua madre; (3) il figlio insegna al proprio figlio solo la lingua ufficiale. Come si vede, nei casi più drastici bastano due generazioni per passare da una lingua completamente viva a una lingua completamente morta. Un caso emblematico è quello dei pigmei, un'etnia molto uniforme e caratterizzata da un punto di vista fisico e culturale (tradizionalmente cacciatori delle grandi foreste del Centro Africa, in mezzo a popoli di agricoltori e allevatori). Tuttavia, non esiste una lingua pigmea (o più lingue pigmee): ogni sottogruppo ha adottato la lingua della popolazione vicina. Come si spiega questo strano fenomeno? Si tratta di un'etnia ritenuta inferiore dai popoli confinanti, in quanto considerati primitivi in base al loro modo preneolitico di produzione economica. Il problema sorge quando si assorbe questo atteggiamento di inferiorità, una manifestazione di auto-stigmatizzazione, come deve essere accaduto ai pigmei.

Assodato che è in atto una massiccia estinzione delle lingue – fenomeno che va avanti da almeno 10.000 anni, ma che sembra aver subito un'accelerazione violenta nell'ultimo secolo – c'è ragione di preoccuparsi? O si tratta di un fenomeno senza alcuna importanza (se non, addirittura, di cui rallegrarsi, visto che elimina alcune 'barriere linguistiche': è un passo necessario verso la 'lingua universale?'). M. Krauss, uno studioso americano delle lingue del Nord America, scriveva già nel 1992:

Ovviamente dovremo ripensare seriamente le nostre priorità, perché la linguistica non passi alla storia come la sola scienza che abbia assistito passivamente alla scomparsa del 90% del proprio campo di indagine.

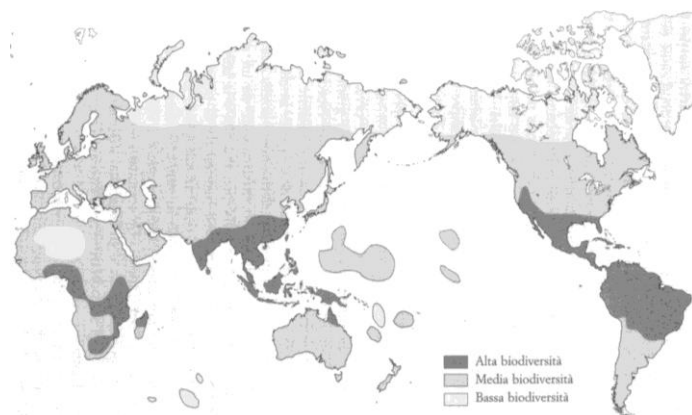
La formulazione di Krauss sottolinea il valore delle lingue come fonte di dati insostituibili per molte scienze dell'Uomo. L'importanza delle lingue sta nel fatto che *non mentono e non dimenticano*. Ogni lingua attuale è l'ultimo anello di una lunghissima catena di trasmissioni verticali (genitori-figli) che risale, con ogni probabilità, all'origine stessa della nostra specie. Attraverso lo studio di una lingua (meglio: di un gruppo di lingue imparentate) è possibile ricostruire la preistoria dei popoli, delle loro migrazioni, delle sedi precedenti, e così via. Inoltre, ogni lingua è il riflesso delle esperienze da parte di ogni gruppo umano dell'ambiente in cui si è sviluppato; rappresenta quindi una peculiare visione del mondo. Basti pensare che le tassonomie presenti in alcune lingue "etiche" sono state utilizzate nelle tassonomie scientifiche, oltre che in medicina, per utilizzare a fini medicinali alcune specie vegetali denominate e utilizzate dagli autoc-

toni. Se una lingua scompare, scompare anche l'esperienza del mondo che le innumerevoli generazioni di persone della comunità che la parla ha impresso in esse. In effetti, ogni lingua è un trattato pratico di etnografia.

Ovviamente, l'importanza per la scienza non è l'unica ragione per essere allarmati dell'estinzione delle lingue – del resto, è assai poco verosimile che i gruppi che parlano lingue a rischio indicherebbero nella scienza la loro preoccupazione primaria. Il problema è che quando si parla di 'cultura', non ci si rende spesso conto che la lingua madre è l'elemento culturale identitario per eccellenza: quando questa va perduta, scompare anche la propria identità culturale. La ragione è che la cultura viene trasmessa attraverso la lingua nativa; se questa va perduta svanisce con essa gran parte delle tradizioni orali, fatta di storie e leggende, di credenze e di canti.



Un'altra ragione per essere allarmati dall'estinzione delle lingue è di natura indiretta: è stato sostenuto da diversi ricercatori che glottodiversità e biodiversità coincidano. La cartina seguente, che mostra la biodiversità presente attualmente nella Terra, mostra in effetti una certa coincidenza con quella della glottodiversità già data sopra:



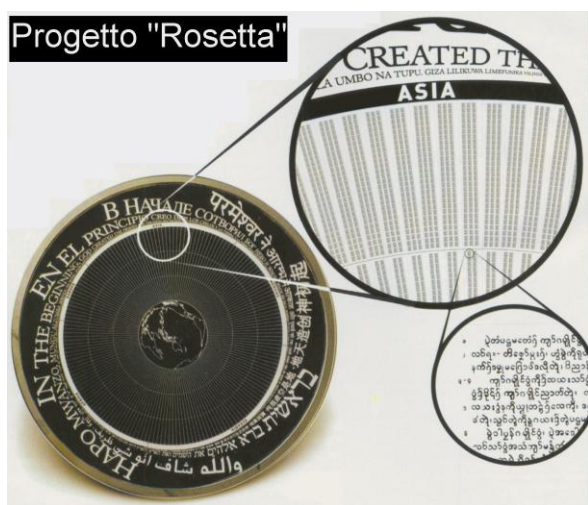
Sebbene la proposta sia stata criticata perché fondata su una sopravvalutazione delle corrispondenze tra i due fattori, resta da chiedersi perché esista un certo grado di coincidenza tra biodiversità e diversità linguistica, dato che si tratta di elementi chiaramente irrelati. La risposta di gran lunga più plausibile è che questi due tipi di diversità condividono lo stesso nemico: 'l'uomo neolitico' (di cui un discendente ancora più potente e pericoloso è 'l'uomo industriale'). Da un lato, in entrambi i casi, la riduzione è stata provocata dall'imposizione, da parte di alcune culture dominanti all'interno del genere umano, di

specie animali e vegetali utili, a discapito delle altre; dall'altro, ha determinato uno status di subalternità di alcuni gruppi umani, e quindi delle loro lingue, rispetto ad altri.

Di conseguenza, diversi studiosi – linguisti, antropologi, sociologi, in particolare – e altre persone sensibili ai problemi della conservazione delle diversità si sono chiesti se non fosse il caso di prendere delle misure per fermare questa catastrofe; se sì, che cosa in concreto si possa fare.

Solo di recente si è trovata la volontà politica e i mezzi finanziari per fare qualcosa al riguardo. Una risposta è stato l'avviamento di alcuni progetti di documentazione delle lingue in via d'estinzione. Uno dei più interessanti è quello denominato 'Rosetta' (dalla famose stele trilingue che permise a Champollion di decifrare i geroglifici egiziani, la cui lingua era allora data per definitivamente perduta), diretto da J. Mason, della *Long Now Foundation* di San Francisco, e finanziato dalla Volkswagen. Ha due obiettivi specifici, già in buona parte realizzati:

(1) il principale è un database digitale di liste di parole per 4000/5000 lingue (per colmare le lacune esistenti nella documentazione il team del Rosetta Project ha allestito da qualche anno un sito web collaborativo, [www.rosetta-project.org](http://www.rosetta-project.org), attraverso il quale studiosi e parlanti di lingue native possono inserire e controllare liste di lessemi, registrazioni audio, grammatiche e altri tipi di documentazione); (2) una "stela" digitale (un dischetto al nichel ad altissima densità del diametro di 76 millimetri, contenente microimmagini leggibili con un microscopio a 1000 ingrandimenti) che contiene 27 pagine di testo commentato e di descrizione in inglese per ciascuna di 1000 lingue, un testo parallelo (il primo capitolo della Genesi), scritto nell'ortografia nativa o traslitterato nel caso in cui la lingua non abbia una scrittura propria, e una mappa (al centro) che indica la localizzazione di ognuna di queste lingue:



La Fondazione conta di produrre questi dischetti in grandissimo numero, protetti da globi di acciaio di protezione, e di distribuirli in tutto il mondo, a beneficio dei nostri lontani posteri.

Tuttavia, benché si tratti di iniziative utili e necessarie, una cosa è conservare la memoria delle lingue ("fossiliz-

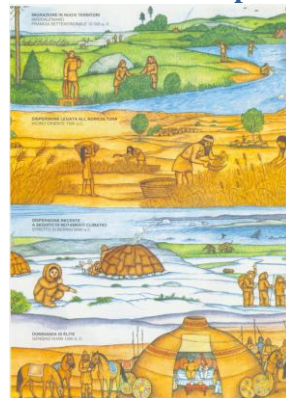
zarle"), altra conservarle in vita. Che cosa si può fare al riguardo? L'idea – a volte sostenuta, a volte addirittura realizzata con qualche esito – di riportare in vita una lingua estinta o di ravvivare una lingua moribonda è poco convincente: una volta che la catena della trasmissione verticale di una lingua è spezzata nulla si può più fare; quello che si ottiene con operazioni di chirurgia linguistica è un oggetto diverso dalla lingua precedentemente parlata. Inoltre, il protezionismo linguistico non sembra aver mai ottenuto successi tangibili: equiparare una lingua parlata da decine o centinaia di milioni di persone – utilizzata nell'istruzione, nella vita burocratica, nella comunicazione tecnico-scientifica, nell'intrattenimento, ecc. – a una lingua parlata da qualche decina di individui, con pochissimi ambiti di uso e scarsissima elaborazione (magari priva di una forma scritta) è un esercizio vano, oggi più che mai, nelle società sempre più dominate dai mass media. Imporre per legge di parlare una lingua piuttosto che un'altra non sembra una soluzione praticabile; inserire lo studio di una lingua morente a scuola può essere di qualche utilità, ma solo a livello di identità culturale: le lingue madri non si imparano a scuola, non più di quanto si diventasse madrelingua di latino attraverso il suo studio. Pertanto, la 'resistenza linguistica' o le 'guerre linguistiche' lasciano il tempo che trovano.

Quella che appare come la soluzione più realistica è stata suggerita, tra gli altri, da J.A. Matisoff, dell'Università di Berkeley:

In definitiva, la risposta al problema dell'estinzione delle lingue è il multilinguismo.

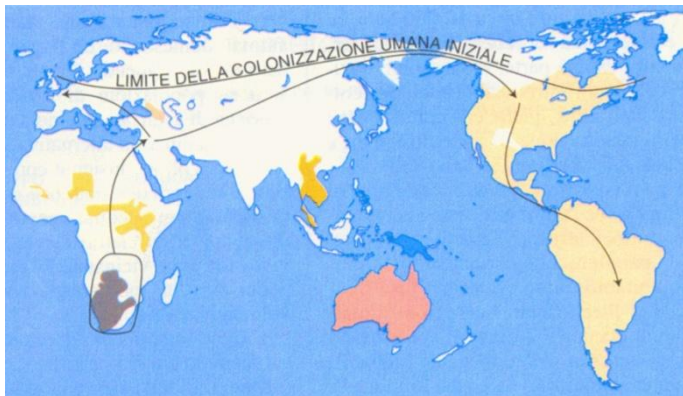
Questo significa che è necessario avviare una seria politica di educazione linguistica che, in primis, sfati i pregiudizi riguardo alla superiorità/inferiorità tra le lingue; inoltre, che sfati il mito del monismo linguistico, i.e. l'idea ottocentesca 'una nazione, un popolo, una lingua' e favorisca il multilinguismo consapevole: quale lingua usare in quale situazione, con chi. Insomma, che ogni individuo ha identità multiple (locale, regionale, nazionale, globale).

## 6. Come si crea la parentela linguistica?



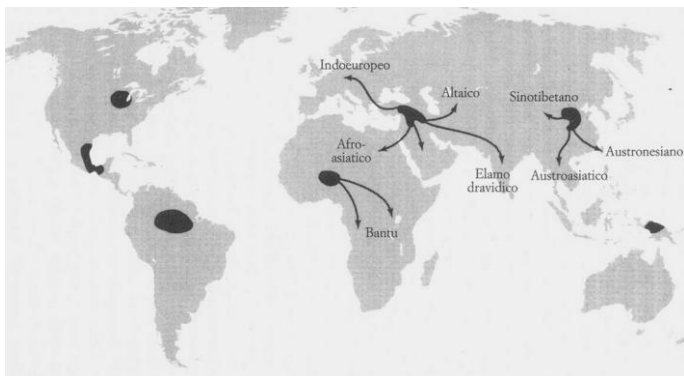
Nel corso di questa trattazione abbiamo sfiorato più volte il problema di come si creino le famiglie linguistiche e la loro fortuna. C. Renfrew, un archeologo britannico specializzato nello studio della cultura antico Indoeuropea,

ha avanzato una proposta che ha il merito di storicizzare un fenomeno che è rimasto generalmente confinato nella pura speculazione, se non nel mito. La prima fase, responsabile della formazione di gran parte della diversità linguistica come la conosciamo oggi, è quella che egli denomina *migrazione in nuovi territori*. Si trattò della diaspora dei gruppi umani nei territori disabitati. Questo fenomeno è intimamente legato al tipo di economia pre-neolitica, basata su caccia/raccolta: quando un territorio riduce le possibilità di nutrire il gruppo umano che ospita, una parte di esso si sposta in nuovi territori. Questo movimento centrifugo, che deve avere accompagnato la nostra specie per quasi tutta la sua storia, ha creato un'enorme frammentazione sociale: si trattava di un numero relativamente esiguo di individui che si spostava in territori immensi. Di conseguenza, questo meccanismo ha creato la gran parte della frammentazione linguistica presente oggi nel mondo. Secondo Renfrew, a questa fase risalirebbero alcune famiglie linguistiche oggi presenti, in base al fatto che si tratta di famiglie che comprendono lingue parlate fino a poche generazioni fa da popoli di cacciatori-raccoglitori:

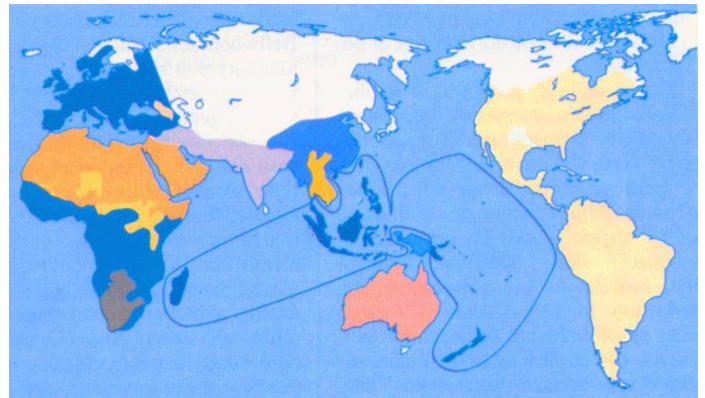


■ KHOISAN  
■ NILO-SAHARIANO  
■ CAUCASICO  
■ AUSTRICO  
■ INDO-PACIFICO  
■ AUSTRALIANO  
■ AMERINDIO

La seconda fase proposta da Renfrew è la *dispersione legata all'agricoltura*, i.e. la rivoluzione neolitica, adottata dapprima da pochissimi gruppi umani, poi diffusasi a macchia d'olio su vasti territori. La cartina seguente illustra i centri di origine dell'agricoltura, con alcune delle famiglie linguistiche associate (o, almeno, le principali):

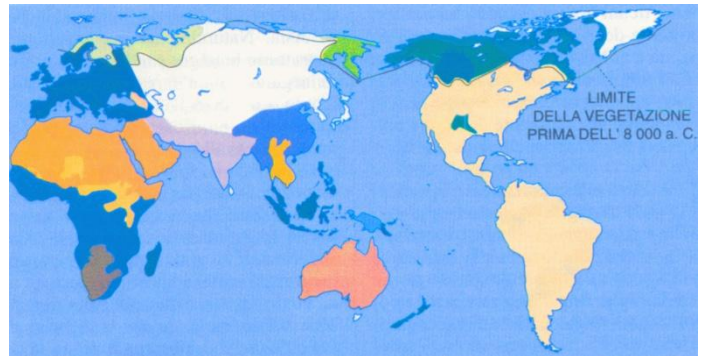


Pertanto, Renfrew ritiene che la situazione linguistica del mondo intorno all'8.000 a.C. (nella prima diffusione del neolitico) fosse come segue:



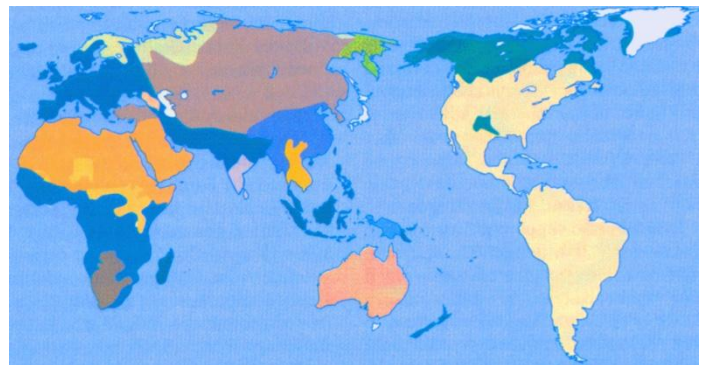
■ NIGER-KORDOFANIANO  
■ CAMITO-SEMÍTICO  
■ INDOEUROPEO  
■ ELAMO-DRAVIDICO  
■ SINO-TIBETANO  
■ AUSTRONESIANO

La terza fase proposta da R. è la *dispersione recente* a seguito di mutamenti climatici, che comportò nuove migrazioni, o "dispersioni", nei territori che si resero abitabili dopo l'ultima glaciazione. La situazione linguistica doveva essere come segue, con la comparsa di famiglie linguistiche artiche:



■ URALICO  
■ CHUKCHI-KAMCHADAL  
■ ESKIMO-ALEUTINO  
■ NA-DENE

L'ultima fase proposta da R., infine, è la *dominanza d'élite*: lo sviluppo di società complesse permise a piccoli gruppi di dominare altre popolazioni e di imporre loro le proprie lingue. La famiglia altaica, comprendente le lingue dei popoli delle steppe, si è diffusa in questo modo (come pure alcune lingue appartenenti a famiglie preesistenti quali l'indoeuropeo e il sino-tibetano):



■ ALTAICO

La proposta di Renfrew è senz'altro molto interessante, se non altro come ipotesi di lavoro, avendo il merito di fon-

dare su base deduttiva interdisciplinare (il modo di produzione economica, il rapporto tra densità demografica ed estensione territoriale, la situazione climatica) la ricostruzione della situazione linguistica del mondo virtualmente "ab origine". Tuttavia, è caratterizzata da una certa confusione tra fasi storiche e processi, che R. mira a unificare, ma a torto: è vero che alcune fasi storiche sono state dominate, se non monopolizzate da un processo, ma non tutte. Ad esempio, la *dominanza d'élite* è un processo che va avanti da c. 12.000 anni, i.e. dalla *rivoluzione neolitica*, fino ai giorni nostri, caratterizzati dalla cosiddetta 'globalizzazione': tra la dominanza di élite dei primi agricoltori-allevatori e l'enorme diffusione dell'inglese contemporaneo non c'è nessuna differenza concettuale. Si tratta di un processo centripeto, che crea grandi concentrazioni sociali, quindi uniformità linguistica (la denominazione di *dispersione*, associata all'agricoltura, è infelice: semmai, si tratta di *diffusione*). La stessa *dispersione legata all'agricoltura* è un fenomeno che ha avuto uno sviluppo molto diseguale: tra la comparsa dell'agricoltura nella Mezzaluna Fertile e nell'America del Nord intercorrono molte migliaia di anni. Allo stesso modo, la *migrazione in nuovi territori* e la *dispersione recente a seguito di mutamenti climatici* sono prodotti dello stesso meccanismo centrifugo: frammentazione sociale, ergo linguistica – da parte di gruppi sociali poco coesi, a causa del loro modo di produzione economica – a seguito della dispersione in enormi territori, via via che si rendevano abitabili. Come è noto, l'abitabilità dei territori per la nostra specie (e specie antenate) è variata molto nel corso del tempo.

Da un punto di vista generale, utilizzando solo un minimo di idealizzazione, si può quindi dividere la storia sociale e linguistica umana in due fasi opposte: quella pre- e quella post-neolitica, ovviamente caratterizzate da un'enorme asimmetria temporale, che hanno creato, rispettivamente, *aumento* e *riduzione* della diversità linguistica della Terra. Con tutta evidenza, il fenomeno della scomparsa delle lingue è cominciato nel neolitico: non sapremo mai quante lingue sono morte nel corso della storia umana. Basti pensare alla situazione linguistica dell'Europa preistorica prima della diffusione delle lingue della famiglia indoeuropea, introdotte da popoli che detenevano le nuove tecniche di produzione economica: a parte il basco e le lingue caucasiche, non sono sopravvissute lingue che ne discendono (qualche traccia è rimasta nella toponomastica e poco più).

Da ciò si evince che per quasi tutta la storia dell'umanità la diversificazione del linguaggio umano deve essere stato più simile a quella della Nuova Guinea che a quella dell'Europa contemporanea.

## 7. Quanto si può risalire indietro nella storia linguistica?

Se il modello di R. fornisce un quadro generale della formazione della diversità/parentela linguistica, non è di nessun aiuto per risalire indietro nella storia dell'umanità in senso strettamente linguistico: appurato che la grandis-

sima parte delle lingue del mondo sono variamente imparentate tra loro, i.e. appartengono a una stessa famiglia (= derivano da una stessa comunità linguistica), fino a che punto è possibile stabilire ulteriori relazioni di parentela tra famiglie linguistiche diverse (magari fino alle estreme conseguenze: la "lingua originaria")?. Il problema è assolutamente aperto e, in una certa misura, empirico: man mano che le conoscenze progredivano si è passati alla individuazione della parentela da quella più prossima (p.e. le lingue romanze, derivate dal latino) a quella meno prossima (p.e. l'Indoeuropeo). Ogni risposta a questa domanda è basata più su opinioni che fatti. Il problema è, in effetti, di natura metodologica: le parentela linguistiche sono state ricostruite tramite un metodo, perfezionato nel corso di due secoli, chiamato *grammatica comparata*, che consiste nel comparare poche lingue a tappeto (almeno per quanto la documentazione lo renda possibile); questo significa, essenzialmente, lessico, morfologia e fonologia. Questo metodo dà molte garanzie che l'affinità tra alcune lingue sia effettivamente il risultato di *eredità* e non di *casualità* (p.e., il latino *dies* 'giorno' è molto simile per forma e significato all'inglese *day*, ma si tratta di una semplice coincidenza: date qualsiasi due lingue del mondo, si troveranno decine di parole che coincideranno in una certa misura per forma e significato) o di *influenza areale posticcia* (p.e., il fatto che in italiano esistano parole come *film* o *sport* non significa che sia strettamente imparentato con l'inglese, ma solo che queste due parole sono prestiti dall'inglese). Tuttavia, il suo limite coincide con la sua virtù: tanto più la *grammatica comparata* ottiene risultati sicuri quanto più le lingue in oggetto sono strettamente imparentate. I maggiori successi di questa linea di ricerca sono state le grandi famiglie che si sono affermate in seguito alla rivoluzione neolitica, Indoeuropeo in primis, quindi 10/12.000 anni fa. Il problema è che le lingue umane hanno una spiccata vocazione al mutamento e, in tutta evidenza, sono state tramandate per molte decine di migliaia di anni. Probabilmente, nessun linguista storico negherebbe che la storia linguistica dell'umanità non è cominciata con famiglie come l'Indoeuropeo; tuttavia, molti di essi si sono mostrati scettici riguardo a ogni tentativo di risalire più indietro. Quindi, il problema non è di principio, ma di metodo: in confronto a quello comparativo ortodosso, gli altri metodi utilizzati non darebbero sufficienti garanzie. Come si noterà, il ragionamento è circolare: il metodo comparativo ortodosso si può applicare solo a situazioni linguistiche relativamente recenti e si esclude la fondatezza di altri metodi comparativi perché non sono come quello comparativo ortodosso.

Un metodo comparativo di natura molto diversa, denominato 'confronto multilaterale', è stato utilizzato da J. Greenberg, dell'Università di Stanford, con le lingue aborigene delle Americhe (in realtà, Greenberg aveva utilizzato molti anni prima lo stesso metodo per l'Africa, con risultati accolti prima con molto scetticismo, poi con consensi pressoché unanimi). La situazione linguistica nativa dell'America, così come era stata ricostruita in base al

metodo della grammatica comparata, è molto complessa, essendo articolata in decine di famiglie diverse, in particolare nella parte subartica del continente:



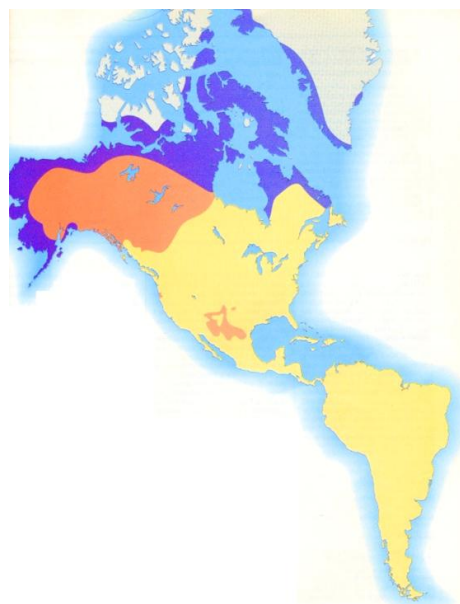
Il confronto multilaterale è sostanzialmente inverso alla grammatica comparata: consiste nel confrontare tutte le lingue di un'area sulla base di una ristretta lista di parole, quelle del lessico fondamentale, i.e. quelle parole di base che ricorrono in tutte le lingue umane (quali parti del corpo e attività connesse allo stesso, rapporti di parentela, ecc.). G. ha proposto circa trecento etimologie. Le tabelle seguenti ne illustrano tre, particolarmente significative, per ragioni che diventeranno chiare più avanti:

GRUPPO	LINGUA		
ALMOSAN	nootka	t'an'a	bambino
KERESIOUAN	yuchi	tane	fratello
PENUTI	totonac	t'ána-t	nipote
HOKA	coahuilteco	t'an-pam	bambino
CENTRO-AMERINDIA	proto-uto-azteco	*tana	figlio-a
CHIBCHA	miskito	tuk-tan	bambino, ragazzo
PAEZ	warrau	dani-	sorella d. madre
ANDINA	aymara	tayna	primogenito
MACROTUCANO	macaca	tani-mai	sorella minore
EQUATORIALE	urubu-kaapor	ta'in	bambino
MACROCARIBICA	pavishana	tane	mio figlio
MACROPANO	lengua	tawin	nipote
MACROGE	tibagi	tog-tan	ragazza
	<b>PROTOAMERIND.</b>	<b>T'AN'A</b>	<b>bambino/a</b>

GRUPPO	LINGUA		
ALMOSAN	yurok	t'in	giovane uomo
KERESIOUAN	mohawk	-tsin	maschio, ragazzo
PENUTI	molale	pnē-t'in	mio fratello maggiore
HOKA	yana	t'inī-si	bambino, figlio/a
CENTRO-AMERINDIA	cuicateco	'dīinó	fratello
CHIBCHA	changuena	sin	fratello
PAEZ	millcayac	tzhəng	figlio
ANDINA	tehuelche	den	fratello
MACROTUCANO	tiquie	ten	figlio
EQUATORIALE	mocochi	tin-gwa	figlio, ragazzo
MACROCARIBICA	yagua	dēnu	bambino
MACROPANO	tacana	u-tse-kwa	nipote
MACROGE	guato	china	fratello più grande
	<b>PROTOAMERIND.</b>	<b>T'IN'A</b>	<b>figlio, fratello, ragazzo</b>

GRUPPO	LINGUA		
ALMOSAN	coeur d'alène	tune	nipote femmina
KERESIOUAN	yuchi	t'one	figlio/a
PENUTI	miwok sierra centrale	tūne-	figlia
HOKA	salinano	a-t'on	sorella piccola
CENTRO-AMERINDIA	taos	-t'ut'ina	sorella piccola
CHIBCHA	lenca	tuntu-rusku	sorella piccola
PAEZ	canapa	t'uh-ki	sorella
ANDINA	tehuelche	thaun	sorella
MACROTUCANO	tiquie	ton	figlia
EQUATORIALE	morotoko	a-tune-sas	ragazza
MACROCARIBICA	nonuya	-tona	sorella
MACROPANO	tacana	-tóna	sorella piccola
MACROGE	piokobyé	a-ton-kä	sorella piccola
	<b>PROTOAMERIND.</b>	<b>T'UN'A</b>	<b>figlia, sorella, ragazza</b>

I risultati sono stati sorprendenti: tutte le lingue delle Americhe sono state raggruppate in tre macrofamiglie, solo le prime due delle quali, quelle più settentrionali: eskimo-aleutino e Na-Dené, coincidono con quelle già accertate col metodo comparativo ortodosso; la grossa novità è rappresentata dalla terza famiglia, che G. ha denominato *Amerindio*, che raggruppa tutte le altre lingue (e tutte le decine di famiglie precedentemente ricostruite), come mostra la cartina:



Dal punto di vista metodologico, va sottolineato che il confronto multilaterale non è poi così diverso dal metodo della grammatica comparata come sembra: l'attendibilità viene ottenuta dal quadro complessivo. Questo significa che, se può essere una semplice coincidenza che due lingue a caso di una macroarea mostrino delle parole simili per forma e contenuto, non può essere una coincidenza che *tutte, o quasi*, le lingue di una macroarea mostrino delle parole in comune. C'è anche da considerare che si comparano parole del lessico fondamentale, ciò che esclude, o rende altamente improbabile, che si tratti di prestiti. C'è un altro aspetto molto interessante nella ricostruzione di G.: come in quelle ottenute tramite il metodo comparativo ortodosso, c'è evidenza di ricostruzione grammaticale, ciò che esclude sia la possibilità di coincidenza sia di prestito, come si evince unendo le tre parole ricostruite che abbiamo mostrato in precedenza:

**PROTOAMERINDIO**

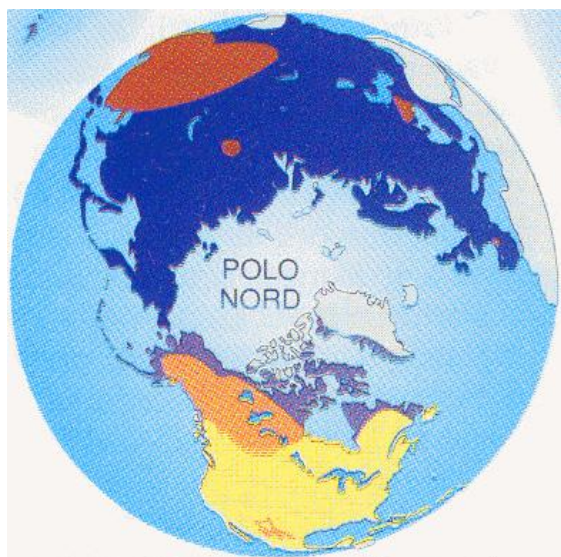
T'AN'A	T'UN'A	T'IN'A
bambino	ragazza	ragazzo
NEUTRO	FEMMINILE	MASCHILE

Come si vede, la vocale tematica (quella interna) muta col genere. Si tratta di un processo morfologico flessivo ben noto, in particolare nella linguistica indoeuropea: l'apofonia. Le chances che si tratti di una semplice coincidenza sono pertanto virtualmente nulle.

Il passo successivo è stato di confrontare le forme derivate dalla ricostruzione delle famiglie americane con quelle delle famiglie del Vecchio Mondo. Anche in questo caso i risultati sono stati impressionanti. La tabella seguente illustra una di queste parole ricostruite, la cui distribuzione areale è enorme:

FAMIGLIA	LINGUA		
AFRO-ASIATICA	proto-AA	*mlg	succhiare, petto, mammella
INDOEUROPEA	proto-IE	*melg	mungere
URALICA	proto-UF	*mälke	petto
DRAVIDICA	tamil	melku	masticare
ESCHIMO-ALEUTINA	yupik centrale	melug	succhiare
AMERINDIA	proto-Amer.	*maliq'a	inghiottire, gola

Il risultato è la scoperta di una nuova superfamiglia, l'Euroasiatico – che comprende diverse famiglie, distribuite attualmente tra Asia e Europa, precedentemente ritenute autonome, come l'Indoeuropeo, l'Afroasiatico, l'Uraloaltaico, Dravidico – che è imparentato direttamente con una delle famiglie americane, l'Eskimo-aleutino. L'euroasiatico è invece solo indirettamente imparentato con l'Amerindio (la famiglia americana più vasta): si tratterebbe di una famiglia sorella, i.e. Euroasiatico e Amerindio sarebbero due ramificazioni di una famiglia linguistica (una comunità linguistica) ancora precedente. In seguito ad altri studi, risulterebbe che la terza famiglia americana, il Na-Dené, sarebbe invece imparentata con lingue molto più occidentali, quali il basco e le lingue caucasiche, con una famiglia asiatica come il sinotibetano e altre lingue asiatiche prima considerate isolate (burushaski, parlato in Pakistan, e lo yeniseiano, parlato in Siberia):



Qualunque sia la verità ultima su queste questioni, resta il fatto che il metodo del confronto multilaterale è l'unica novità nel settore della ricerca della ricostruzione della diversità linguistica antica. Il metodo comparativo ortodosso è di fatto incapace di risalire oltre *qualche migliaio* di anni; il metodo del confronto multilaterale si è dimostrato in grado di risalire indietro di *qualche decina di migliaia* di anni, operando una drastica riduzione delle famiglie linguistiche del mondo, un risultato certamente apprezzabile. Tuttavia, al momento nemmeno questo metodo sembra in grado di risalire più indietro nella storia linguistica dell'umanità: rimangono una serie di famiglie linguistiche in Eurasia e in Africa irriducibilmente diverse. Sarà mai possibile andare oltre? Per il momento, non è possibile dare una risposta fondata a questa domanda. A dire il vero, ci sono stati dei tentativi in questa direzione da parte di M. Ruhlen, uno dei più stretti collaboratori di Greenberg, che ha proposto una trentina di 'etimologie globali', i.e. parole ricostruite che sarebbero comuni a tutte le famiglie del mondo, persino a quelle africane, che sembrano tra tutte quelle più irriducibilmente diverse, vista la maggiore antichità della loro dispersione. Anche in questo caso, alcune corrispondenze sono impressionanti, e non si lasciano liquidare facilmente come coincidenze. Il problema di questa operazione, tuttavia, è che risulta fine a se stessa: spezza solo, al massimo, una lancia a favore della teoria monogenetica del linguaggio (i.e. tutte le lingue derivano da una lingua primigenia) o, meglio, che tutte le lingue sopravvissute derivano da una stessa lingua, ma nulla più. In realtà, non ci dice nulla della storia linguistica dell'umanità arcaica, ma solo di quella proto-linguistica. In altre parole, è astorica, né si vede per ora la possibilità di associarla a dati extralinguistici come quelli di Renfrew.

**8. Gruppi etnici e lingue**

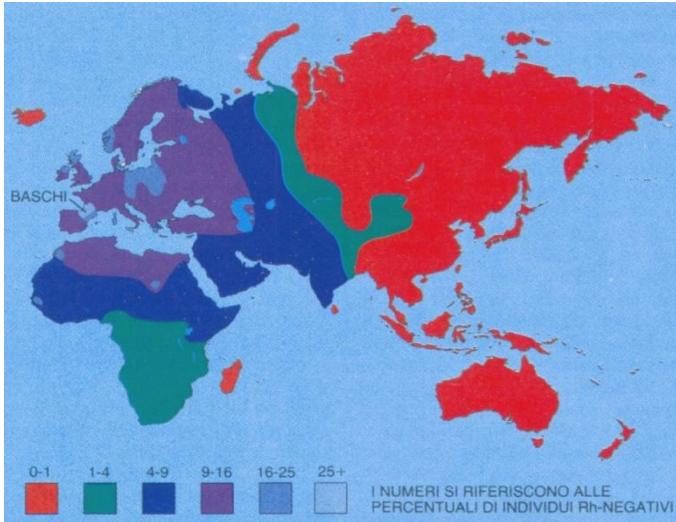


Uno sviluppo più interessante al problema della storia linguistica dell'umanità è rappresentato dagli approcci interdisciplinari. Una ricerca particolarmente interessante in questo ambito è quella di L. Cavalli Sforza, un genetista italiano della Stanford University, che ha lavorato sulle differenze nell'ambito del genoma umano.

Una premessa è d'obbligo per evitare pericolosi fraintendimenti. I gruppi umani ottenuti su base genetica non equivalgono alle 'razze' nell'uso comune del termine. Il concetto di razza, per lo più individuato a partire dal colore della pelle e di altre caratteristiche fisiche eventenziali, è un mito culturale. Ad esempio, non esiste una razza 'nera/negra' (o come altro viene denominata): esistono sulla Terra molte popolazioni con la pelle molto pigmentata che, però, non hanno nulla in comune in esclusiva dal punto di vista genetico. Un aborigeno australiano o papuano o melanesiano o un indiano del sud, tutti normalmente definiti 'neri' hanno molto più in comune

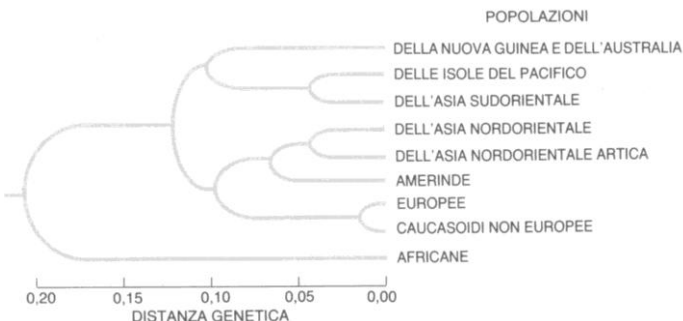
geneticamente con un europeo che con un africano. Una razza 'nera' non esiste nemmeno in Africa: tra le popolazioni africane esistono differenze genetiche notevoli, come ci si aspetta considerando che l'Africa è il continente dove la dispersione umana è cominciata.

La tipica configurazione nella diversità genetica intra-umana è illustrato dal fattore sanguigno Rh-negativo, che si presenta differenziato nella mappa seguente:



Ci sono due aspetti importanti: il primo è che si tratta di una *caratteristica non evenemenziale*, i.e. un tratto genetico che non si manifesta esteriormente, ergo non si presta a essere oggetto di pregiudizi razzisti; il secondo è che è *tipicamente graduale e di ordine quantitativo*: non si tratta di tutto o niente, ma di percentuale variabile, e individui dello stesso gruppo (persino fratelli) possono differire per questa caratteristica. Come si vede, la percentuale dell'incidenza del fattore Rh-negativo decresce allontanandosi da un epicentro costituito dai Paesi Baschi e risulta pertanto quasi assente in Oriente (e quindi nelle Americhe, sebbene esse non siano mostrate nella cartina). È evidente che appaiare questi dati scientifici al concetto comune di 'razza' è impossibile.

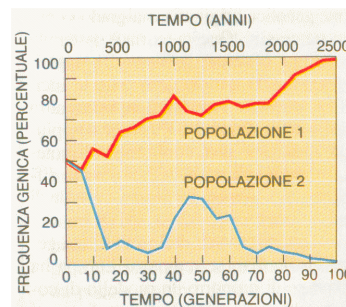
L'albero genealogico ottenuto dai dati genetici di C. è il seguente:



Come si vede, la più grande bipartizione umana è tra popolazioni africane e non africane, risultato del fatto che un gruppo (o più gruppi omogenei) lasciò l'Africa per colonizzare il resto della Terra qualche decina di migliaia di anni fa. L'altra grande bipartizione è tra euroasiatici "set-

tentrionali" e "meridionali", frutto di un'altra grande diaspore.

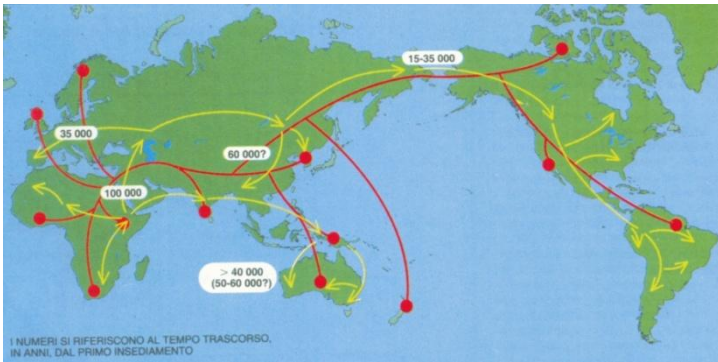
Un altro elemento interessante utilizzato nello studio di C. è la datazione assoluta. Gli alberi genealogici (come quelli linguistici) forniscono solo datazioni relative, i.e. una diramazione ne precede un'altra, e così via, esattamente come un padre precede un figlio, ma non dicono nulla riguardo al tempo in cui la divisione si è effettivamente verificata; sono quindi astorici. Tuttavia, i genetisti hanno messo a punto un "orologio genetico", che è in grado di misurare il tempo con buona approssimazione. Si tratta del DNA mitocondriale, che è particolarmente adatto allo scopo per due ragioni: (1) viene trasmesso solo per via materna (allo stesso scopo è stato a volte usato anche il cromosoma Y, trasmesso solo per via paterna); (2) non è soggetto a selezione ambientale in quanto non produce manifestazioni evenemenziali o funzionali (la maggior parte dei geni sembrano in effetti di questo tipo), ma è connesso alla generazione di energia cellulare. Quindi, muta con molta regolarità e consente di calcolare approssimativamente quanti anni sono intercorsi tra un ramo e l'altro dell'albero genealogico (i.e. si può dire quando due gruppi si sono separati). Per quanto riguarda le lingue, non esiste purtroppo nulla di comparabile. Una proposta, il metodo 'glottocronologico', è stata avanzata negli anni '50 da M. Swadesh, un linguista americano, basata sull'assunto che il mutamento nel lessico fondamentale di una lingua (da lui individuato in 215 parole universali) sia costante nel tempo. Il metodo si è dimostrato inaffidabile una volta applicato a lingue di cui si conosce bene la storia: è vero che il lessico elementare è relativamente stabile nel tempo, ma ciò non equivale a mutare costantemente. In effetti, l'assunto sembra infondato: un mutamento del lessico nel senso di sostituzione di parole non è un processo puramente interno, ma spesso il risultato del contatto linguistico. Al contrario, un orologio linguistico dovrebbe essere simile al DNA mitocondriale: un elemento che DEVE mutare per qualche ragione interna e lo fa in modo costante in quanto non soggetto a pressioni esterne. Un elemento del genere non è stato ancora trovato. In conclusione, non esiste (ancora?) un orologio linguistico.



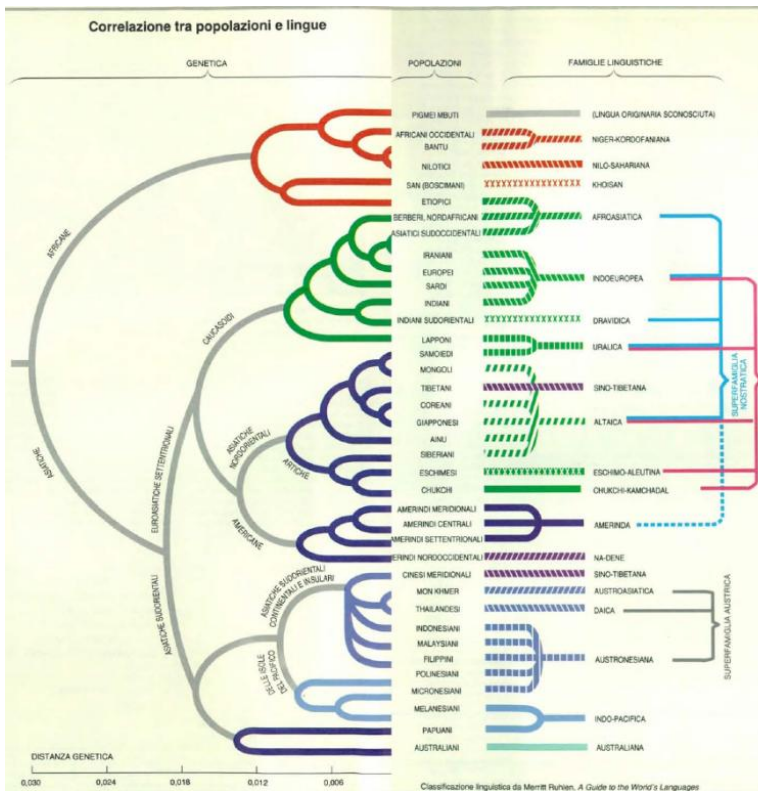
Un modello della deriva genetica può essere ottenuto al computer. Quando due sottogruppi di una popolazione si separano, i geni di ognuno di essi hanno frequenze in un primo momento uguali, ma nel tempo si evolvono in direzioni diverse, come il grafico a lato esemplifica.



Sulla base delle mutazioni genetiche considerate e della mutazione del DNA mitocondriale, la storia della diaspora umana è rappresentabile come segue:



Dopo aver raccolto campioni genetici di tutte le popolazioni autoctone conosciute e averne studiato le peculiarità geniche, C. ha incrociato i dati risultanti con quelli basati sulle famiglie linguistiche. I risultati sono stati sorprendenti: l'albero genealogico delle popolazioni umane corrisponde in buona misura all'albero delle lingue parlate nel mondo; entrambi presuppongono migrazioni che, a quanto indicano i dati biologici, avrebbero avuto origine in Africa. La figura seguente illustra:



La sorpresa riguardo alla coincidenza tra famiglie linguistiche e famiglie genetiche deriva dal fatto che geni e lingue non hanno di per sé nulla in comune: come si sa da almeno duecento anni, le lingue non vengono trasmesse dai genitori ai figli come i geni. Qualsiasi persona, a prescindere dal proprio patrimonio genetico, può imparare una qualsiasi lingua la mondo: non si nasce con la capacità di parlare una lingua né con una particolare predisposizione per una lingua specifica; la lingua madre viene acquisita in base all'ambiente in cui si cresce. Del resto, an-

che i risultati della ricerca di C. mostrano dei casi di spaiamento eclatanti: gli etiopi hanno un DNA africano, ma appartengono linguisticamente a una famiglia asiatico-nordafricana, l'Afroasiatico; i già citati pigmei hanno caratteristiche genetiche esclusive in comune (i.e. sono una popolazione individuale), ma non hanno un'identità linguistica comune.

Come si spiega allora una coincidenza di questa entità? La risposta è che diversificazione etnica e linguistica presentano molte analogie, nonostante le radicali differenze intrinseche. Una è la *vocazione al mutamento*. Mutazione genica e mutamento linguistico, pur non condividendo alcun fattore causale, operano in modo simile: entrambi sono ineluttabili, pervasivi e sostanzialmente casuali (il mutamento linguistico non è del tutto casuale, ma rimane il fatto che non è possibile prevedere come una lingua si evolverà nel lungo periodo; inoltre, se una comunità linguistica si divide in due sottocomunità queste si evolvono tipicamente in modo diverso). L'altra analogia è il fattore sociale, che determina se una mutazione/un mutamento crea o meno diversità genetica e linguistica: se una innovazione, genica o linguistica, si trasmette uniformemente in un'intera comunità, il risultato sarà il cambiamento, ma non la differenziazione. Affinché mutazione o mutamento producano differenziazione è necessaria frammentazione sociale. Questo spiega la coincidenza – in linea di principio, non di necessità – tra gruppi umani e lingue: la differenziazione nella popolazione umana e nelle lingue ha seguito il corso della diaspora umana. Si ricordi, ancora una volta, che le lingue non esistono; esistono le comunità linguistiche, che, fino a prova contraria, coincidono con le comunità umane tout court. Il concetto di comunità, etnica o linguistica, è fondamentale perché, in entrambi i casi, geni e lingue si trasmettono *verticalmente*, i.e. di generazione in generazione (sebbene con le dovute differenze).

Lo studio di Cavalli Sforza ha ricadute sul problema della ricostruzione linguistica arcaica: dato che, come abbiamo visto, esiste un'alta incidenza statistica nella corrispondenza tra popolazioni e lingue, si possono utilizzare i dati genetici per ricostruire, con alta probabilità di affidabilità, anche la storia linguistica dell'umanità. Quali ipotesi linguistiche possiamo trarre dai dati genetici? La prima è che la grande divisione è tra lingue africane (sub-sahariane) e le altre: almeno in linea di principio, si potrebbe arrivare a ricostruire la lingua della popolazione che alcune decine di migliaia di anni fa lasciò il continente africano per colonizzare il resto del mondo. La seconda ipotesi è che l'altra grande divisione è quella tra le lingue euroasiatiche settentrionali e quelle meridionali. La terza ipotesi è che le famiglie africane in senso stretto (sub-sahariane) sono tra loro non meno divergenti le une con le altre di quanto ciascuna di esse lo sia rispetto all'"Euroasiatico" (inteso come tutte le lingue supersahariane); detto in altro modo, qualsiasi coppia di lingue extra africane sono tra loro più strettamente imparentate di quanto lo siano due lingue appartenenti a due diverse famiglie africane.

Ovviamente, questa ricostruzione basata su dati genetici è molto generica, allo stato attuale della ricerca, ma ha il pregio di fornire un quadro generale che generi delle ipotesi di lavoro molto interessanti.

Per quanto riguarda la ricostruzione della parentela linguistica delle Americhe, in particolare, i dati genetici hanno confermato l'ipotesi di Greenberg: come mostra la mappa delle migrazioni umane data sopra, le tre macrofamiglie americane corrispondono ad altrettante ondate migratorie dall'Asia Nordorientale, la prima delle quali è stata quella della popolazione amerindiana, seguita dalla Na-Dené e infine dalla Eskimo-aleutina (che del resto è sparsa anche sul continente asiatico). Anche i dati archeologici (relativi alle caratteristiche dentarie) sono in linea con queste conclusioni. Un caveat è qui necessario: da studi più recenti, c'è ragione di credere che le ondate migratorie in America siano state più di tre; tuttavia, sono tre le ondate migratorie di successo. Inoltre, non è detto, né probabile, che le migrazioni siano state nette come la ricostruzione sembra implicare: è possibile che le varie ondate migratorie si siano protratte per un certo lasso di tempo interessando popolazioni estremo nordorientali geneticamente e linguisticamente molti simili ma non identiche. Questo significherebbe che la differenziazione genetica e linguistica di una popolazione poteva, in una certa misura, preesistere alla colonizzazione dell'America. D'altra parte, non si può chiedere a una ricostruzione di questa antichità più dettaglio di quanto non sia in grado di dare: si tratta semplicemente di un costrutto sensibilmente idealizzato.

## 9. Conclusioni

Le conclusioni che si possono trarre da questa lunga discussione sono le seguenti.

Le lingue, così come vengono concepite dal senso comune, non esistono in natura; quello che esiste è la diversità linguistica, pervasiva e normalmente graduale, che caratterizza ogni comunità linguistica rispetto alle altre e anche, in una certa misura, al proprio interno (e persino all'interno dei singoli parlanti). Quello di lingua (come quello di 'razza') è in parte un epifenomeno, in parte una convenzione, in parte un mito cognitivo-culturale.

Di conseguenza, non c'è nulla di discreto da contare. Pertanto, rispondere alla domanda, ingannevolmente semplice, riguardo al *numero delle lingue del mondo* è in linea di principio impossibile, se non al prezzo di una notevole dose di arbitrarietà. Su di un piano convenzionale, la risposta dipenderà dal metro di misura adottato, i.e. il grado di diversità ('distanziamento') considerato pertinente. Anche in questo caso, però, non ci si può aspettare risultati omogenei, perché nessuno ha il misurino della diversità linguistica. Inoltre, spesso altri fattori giocano un ruolo nella decisione: politici, tradizionali, sociali (il grado di 'elaborazione' di una lingua, in primis). Pertanto, dato che ricercatori diversi produrranno descrizioni diverse e che le descrizioni globali sono la somma di mol-

tissime descrizioni parziali di ricercatori diversi, il risultato non sarà nemmeno omogeneo. Tutto ciò spiega il notevole divario nelle stime prodotte, che sono un utile riferimento pratico, a patto di non prenderle troppo sul serio.

Un'altra caratteristica delle lingue è l'enorme asimmetria nel numero delle persone che le parlano: si va da lingue parlate da centinaia di milioni parlanti a lingue parlate da così pochi individui che non si sa se siano ancora viventi. In effetti, la gran parte delle lingue, quasi il 90%, è parlata da qualche decina di migliaia di persone al massimo. Le grandi lingue nazionali parlate da milioni di persone, le quali si sono espanse a spese delle altre, sono l'eccezione umana, non la regola. Questo vuol dire che quasi il 90% delle lingue del mondo è a rischio di estinzione, di cui più della metà di queste, secondo alcuni, entro il secolo attuale.

Per ovviare al problema della scomparsa del patrimonio linguistico dell'umanità, che rappresenta un danno indicibile per gli studi dell'Uomo, sono stati avviati alcuni progetti di conservazione della documentazione linguistica (lessico, descrizioni grammaticali, scrittura, ecc.), da trasmettere a futura memoria. Per contrastare la tendenza alla scomparsa delle lingue, che va avanti dal neolitico, invece, non sembrano efficaci le politiche di conservazionismo linguistico: l'unica strategia realistica sembra il plurilinguismo, che non rifiuta, ma nemmeno accetta passivamente, le asimmetrie sociali tra lingue.

Anche la distribuzione della diversità linguistica per aree della Terra è assai asimmetrica, in dipendenza della storia sociale e politica, a loro volta in genere conseguenza delle caratteristiche geografiche: minima in Europa, massima (in rapporto al numero dei parlanti) in Oceania e Africa.

Un'altra asimmetria che caratterizza le lingue del mondo è la relazione di parentela: lingue diverse sono imparentate in modo diverso con le altre lingue. La grammatica comparata, la scuola che ha dominato la linguistica del XIX secolo, ha specificato in una certa misura queste parentele ricostruendo molte famiglie linguistiche, alberi genealogici che caratterizzano le lingue parlate attualmente e quelle attestate anticamente. A questo fine, ha messo a punto un metodo di comparazione molto analitico e affidabile, che consiste nel comparare a tappeto poche lingue, per discriminare tra elementi ereditati ed elementi semplicemente presi in prestito da lingue vicine o semplici coincidenze. Il più grande successo di questo metodo è stata la ricostruzione della famiglia Indoeuropea, che rimanda a una comunità linguistica risalente a diverse migliaia di anni fa.

La parentela linguistica è il risultato di due processi inversi. Il primo è *centrifugo*, provocato dalla dispersione sociale in nuovi territori (o in territori resi, o ri-resi, accessibili dalle mutate condizioni climatiche), che crea differenziazione linguistica e che, avendo caratterizzato quasi tutta la storia dell'umanità, ha creato la diversificazione linguistica presente oggi sulla Terra. L'altro processo è *centripeto* ed è cominciato sulla Terra 10/12.000 an-

ni fa, in seguito alla rivoluzione neolitica, che ha conferito ad alcune popolazioni, quindi comunità linguistiche, un vantaggio economico, quindi demografico, considerevole; questo processo crea uniformità linguistica, imponendo una lingua a spese di altre. Le grandi famiglie linguistiche ricostruite nel XIX secolo, come l'Indoeuropea o la Semitica (poi riconosciuta come un membro dell'Afroasiatico), sono il risultato di questo processo.

La globalizzazione attuale non è che l'anello più recente (e più evidente) di una lunga catena che si snoda per molti millenni e che si basa su delle asimmetrie di natura essenzialmente economica.

La grammatica comparata è il metodo più affidabile per ricostruire stadi linguistici relativamente vicini, ma è inutile per quelli più remoti. A questo fine è stato proposto un metodo, il confronto multilaterale, che consiste nel comparare tutte le lingue di un'area in base a pochissimi elementi del lessico. Questo ha permesso di risalire a stadi linguistici considerevolmente più remoti di quelli ricostruiti dalla grammatica comparata (qualche decina di migliaia di anni fa), sebbene con un livello di dettaglio incomparabilmente inferiore. Tuttavia, siamo ben lontani dalla possibilità di accedere a stadi linguistici veramente remoti (a cui forse non accederemo mai), a dispetto delle 'etimologie globali' proposte, che ci aiutano ben poco a ricostruire la storia linguistica, e quindi sociale, dell'umanità.

Un aiuto di natura molto diversa, a questo proposito, proviene dai dati genetici: lingue e geni, i.e. comunità linguistiche e popolazioni mostrano una grande coincidenza statistica (in base ad analogie indirette), sebbene non siano minimamente correlati direttamente. Il fatto che sia possibile – a differenza delle lingue, almeno per il momento – sia ricostruire la storia genica dell'umanità *ab ovo* sia datarla con buona approssimazione fornisce un aiuto indiretto alla ricostruzione della diversificazione linguistica dell'umanità. Ovviamente, si tratta di generalizzazioni che possono fornire delle linee guida per la ricostruzione linguistica, non di ricostruzione linguistica in senso stretto.