

PUBBLICAZIONE DELLA  
SOCIETÀ DI LINGUISTICA ITALIANA  
43

SLI  
SOCIETÀ DI LINGUISTICA ITALIANA

«DATI EMPIRICI E TEORIE  
LINGUISTICHE»

ATTI DEL XXXIII CONGRESSO INTERNAZIONALE DI STUDI  
DELLA SOCIETÀ DI LINGUISTICA ITALIANA  
Napoli, 28-30 ottobre 1999

a cura di  
FEDERICO ALBANO LEONI,  
ELEONORA STENTA KROSBAKKEN,

ROSANNA SORNICOLA,  
CAROLINA STROMBOLI

BULZONI ROMA 2001

PETER KOCH - PATRICK O. STEINKRÜGER  
(Università di Tubinga)

## Poligenesi lessicale e dati «empirici»\*

### 1. PRESENTAZIONE DEL PROGETTO

Nel gennaio 1999 si è creato all'Università di Tubinga un Centro di Ricerca Interdisciplinare *Strutture di dati linguistici* (SFB 441), finanziato dalla Deutsche Forschungsgemeinschaft (Società Tedesca della Ricerca). Si tratta di 13 progetti che costituiscono tre gruppi principali all'interno del progetto complessivo. Ciò che accomuna i singoli progetti è la problematica dei dati empirici, cioè un complesso di questioni che, fino a qualche anno fa, non aveva alcun rilievo nella ricerca linguistica. Benché i dati costituiscano il fondamento della ricerca, i progetti che fanno parte del Centro di Ricerca Interdisciplinare in questione mostrano una grande diversità (vedi Fig. 1).

Fra questi progetti, il nostro (B6: *Mutamento lessicale - poligenesi - costanti cognitive*) rappresenta un caso abbastanza particolare per quanto riguarda la tipologia dei dati. Si tratta, infatti, dell'unico progetto i cui dati si basano esclusivamente sulla lessicografia. Più esattamente i nostri dati consistono di informazioni lessicografiche di tipo etimologico o comparativo. Sono questi i dati che nel futuro costituiranno il materiale di base della nostra banca-dati.

Questa maniera empirica di procedere si spiega col fatto che il progetto riprende, in chiave cognitiva, l'approccio onomasiologico magistralmente esemplificato da Carlo Tagliavini (1949) nel suo lavoro sulla pupilla, ma anche da Charles Buck (1949) nel suo dizionario concettuale delle lingue indoeuropee. Come ha indicato Paola Benincà (1994: 696) nel suo cenno sulla linguistica italiana, esistono precursori dello studio onomasiologico come Oehl (1922) ed è anche possibile trovare una dettagliata bibliografia in Quadri (1952), rivista, poi, con speciale riguardo all'italiano, da Corrà (1981). Sono da menzionare, nell'ambito delle lingue europee, anche il *Dictionnaire onomasiologique des langues romanes* di Vernay (1991 ss.) e il dizionario, purtroppo non portato a termine, di Schröpfer (1979 ss.). Anche il nuovo dizionario comparativo-etimologico delle lingue austronesi, edito da Darrell Tryon (1995), possiede una struttura onomasiologica, cioè i lessemi spiegati seguono i concetti che esprimono, invece di essere classificati secondo la loro provenienza etimologica proto-austronese. Il nostro progetto si occupa

\* Ringraziamo Silvia Mongili della revisione stilistica di questo contributo.

della denominazione delle parti del corpo umano, rintracciando eventualmente gli universali cognitivi responsabili di ogni denominazione. Procedendo in questo modo, potrebbe essere verificata la tesi sulla 'poligenesi lessicale' (cfr. Koch 1997, 1999a e c, in stampa; Blank 1998): in altre parole si tratta di chiarire se certe lingue, senza mai essere venute a contatto tra di loro, arrivino ai medesimi risultati lessicali per designare gli stessi concetti (nel nostro caso, le parti del corpo umano). Il risultato del nostro lavoro porterà alla costituzione di una banca-dati tramite la quale si potrà risalire comodamente alla storia ed al legame cognitivo (p. es. metonimico o metaforico) di qualsiasi denominazione nelle diverse lingue. Tali informazioni saranno disponibili su CD-ROM e sulla rete.

Detto questo, vorremmo ora illustrare due importanti aspetti riguardanti la complessità dei dati empirici, vale a dire i campioni delle lingue esaminate e la qualità dei nostri dati.

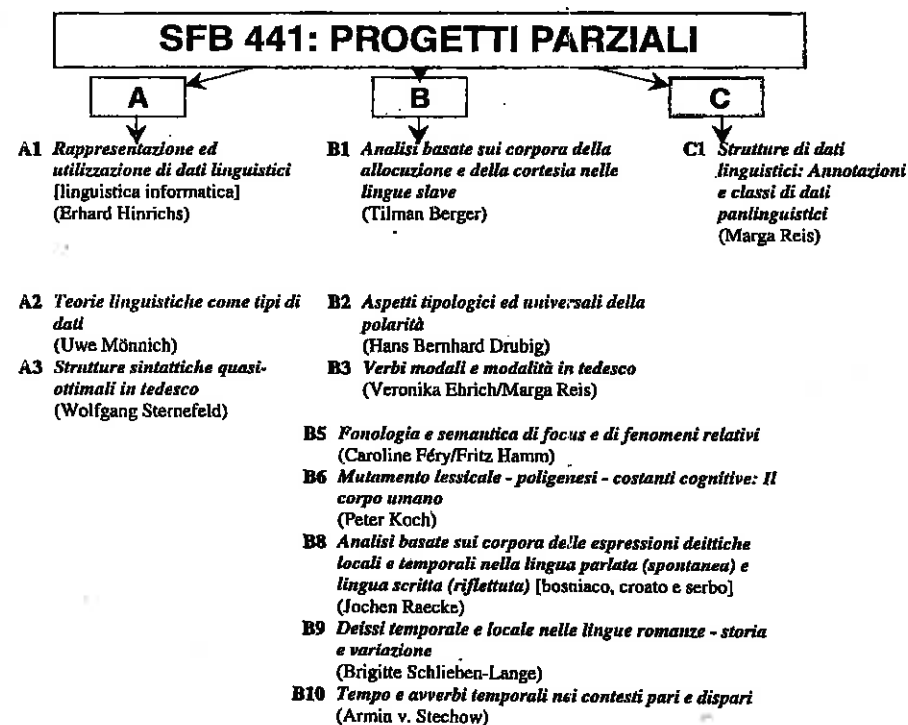


FIGURA 1

## 2. IL PROBLEMA DEL CAMPIONAMENTO

In primo luogo, è necessario elaborare un campione di lingue europee ed extraeuropee sufficientemente ampio e diversificato per poter giungere a dei risultati significativi e rappresentativi, innanzitutto per quanto riguarda un eventuale carattere poligenetico dei fenomeni in esame.

La ricerca riguardante la diacronia lessicale comporta qualche problema specifico. Innanzitutto ci troviamo di fronte ad un problema puramente tecnico: la composizione del campione è strettamente legata all'esistenza ed all'accessibilità di dizionari etimologici e comparativi. Esistono, però, anche delle forti limitazioni dovute alla mancanza, in seno a certe lingue, di una codificazione scritta. Per questo, la maggior parte delle lingue ben documentate appartiene alla famiglia indoeuropea o a quella afro-asiatica. Non parliamo del caso complesso delle scritture ideografiche che si sono manifestate soprattutto nella sfera della cultura cinese.

Se finora abbiamo parlato delle difficoltà inerenti a circostanze esterne, è giunto il momento di esaminare un aspetto concernente un punto di vista teorico. Se per la tipologia grammaticale è fondamentale disporre di un campione di lingue appartenenti a famiglie differenti e di una grande varietà di strutture grammaticali, questo non risulta essere, invece, per forza valido per una ricerca diacronica del lessico. Occorrerà pertanto verificare se i criteri utilizzati nell'ambito di una tipologia grammaticale (p. es. Nichols 1992; Rijkhoff *et alii* 1993; Haspelmath 1997: 15-20) possano essere validi anche nell'ambito di una lessicologia diacronica comparativa. Gli esempi che seguono rendono palese che un'affinità tipologica o genetica non è legata obbligatoriamente ad uno stesso sviluppo lessicale.

Cominciamo con un esempio (cfr. tab. 1), in cui quattro lingue di una stessa discendenza linguistica mostrano, per lo stesso concetto, denominazioni completamente diverse dal punto di vista cognitivo. Il caso tratta della denominazione della bocca umana nelle diverse lingue slave. Evidentemente tutte queste lingue appartengono alla stessa famiglia linguistica e sono simili dal punto di vista tipologico-grammaticale. Ciononostante per la denominazione della bocca umana troviamo soluzioni lessicali divergenti dal punto di vista cognitivo.

Continuiamo con un esempio esattamente opposto, in cui due lingue tipologicamente diversissime mostrano, per lo stesso concetto, due denominazioni completamente analoghe dal punto di vista cognitivo (cfr. tab. 2).

Tutto questo significa, allora, che i criteri da utilizzare per una scelta delle lingue in relazione alla storia dei lessemi, non sono gli stessi necessari, invece, alla tipologia genetica o grammaticale. Bisogna, per esempio, considerare i contatti culturali, i tabù o la particolarità di un lessema all'interno di un contesto religioso ecc.

TABELLA 1  
SLAVO BOCCA

concetto d'arrivo	dato odierno	meccanismi cognitivi <sup>1</sup>		dato storico / polisemia odierna	concetto di partenza
BOCCA	russo <i>rot</i>	similarità (forma, animale → uomo)		slavo orientale antico <i>rbfb</i> 'punta, becco'	BECCO
BOCCA	sloveno <i>ŭsta</i> (pl.)	contiguità (partonomia)		protoslavo * <i>usta</i> (pl.) 'labbra' < ieg. * <i>usta</i> : 'labbro'	LABBRA
BOCCA	ceco <i>pusa</i>	contiguità (funzionalità)		anche 'baccio' < tedesco bavarese <i>puss</i> 'baccio'	BACCIO
BOCCA	casciubo <i>gęba</i>	contiguità (partonomia)	labbro dati comparativi: alto sorabico <i>huba</i> , russo <i>guba</i> 'labbro'	similitudine (qualità) slavo antico <i>goba</i> 'fungo'	FUNGO

(cfr. Machek 1957; Pokorny 1959/69; Vasmer 1953; Trubačev 1979)

TABELLA 2  
ARABO ED UNGHERESE PUPILLA

concetto d'arrivo	dato odierno	meccanismi cognitivi	dato storico / formazione delle parole	concetti di partenza
PUPILLA	arab. ' <i>aswad</i> al' <i>ain</i> 'pupilla'	appartenenza ad una classe tassonomica + contiguità (partonomia)	' <i>aswad</i> 'nero' + ' <i>ain</i> 'occhio'	OGGETTO NERO + OCCHIO
PUPILLA	ungh. transilv. <i>szemfeketéje</i> 'pupilla'	appartenenza ad una classe tassonomica + contiguità (partonomia)	<i>szem</i> 'occhio' + <i>fekete</i> 'nero'	OGGETTO NERO + OCCHIO

(cfr. Tagliavini 1949: 238 s.)

<sup>1</sup> Per l'applicazione di questi meccanismi cognitivi universali al dominio del corpo umano, cfr. Andersen (1978); Schladt (1997); Wilkins (1996).

## 3. ACCESSIBILITÀ E GRADO DI ELABORAZIONE DEI DATI

Nel quadro di un Centro di Ricerca Interdisciplinare consacrato alla problematica dei dati linguistici, al nostro progetto spetta un posto un po' particolare in quanto i suoi «dati» non hanno un carattere «immediato», direttamente accessibile. Infatti, essi vengono ricavati dai dizionari etimologici e sincronici delle lingue esaminate: dizionari che non contengono, però – o almeno in gran parte non contengono – dei dati primari, ma delle informazioni derivate da essi, elaborate e talvolta interpolate.

Per impostare questo problema, abbiamo bisogno di un modello teorico dell'elaborazione e dell'accessibilità dei dati. A nostro avviso bisogna sin dall'inizio distinguere due dimensioni del problema:

- 1) il grado di accessibilità dei dati in quanto tali. I dati appoggiati da un contesto linguistico o extralinguistico sono più accessibili dei dati fuori contesto. I dati grafici, in quanto visuali, sono più accessibili dei dati fonici, in quanto acustici (a prima vista, questa affermazione sembra sorprendente, ma risulta perentoria, se si esaminano da vicino la natura dei due mezzi di comunicazione e la loro interazione con la percezione umana). I dati odierni sono più accessibili dei dati storici (non dimentichiamo che un dato storico prima dell'invenzione della tecnica fonografica è sempre un dato grafico). I dati sincronici sono più accessibili dei dati diacronici (perché i processi diacronici si devono sempre interpolare tra stati di fatto sincronici: v. *infra*). Si noti che non vanno identificati 'odierno' con 'sincronico' da una parte, e 'storico' con 'diacronico' dall'altra, perché esistono anche dei dati storici sincronici.
- 2) il grado di elaborazione dei dati. I dati primari che si trovano nei testi o addirittura negli atti comunicativi individuali devono essere elaborati dai linguisti per acquisire una quanto minima pertinenza linguistica. Devono essere assegnati a categorie fonetico-fonologiche, grammaticali e lessicali, devono essere interpretati dal punto di vista semantico o pragmatico ecc. In questo modo saranno trasformati in dati secondari che a loro volta possono essere trasformati, a livelli più astratti, in dati terziari ecc. ecc. Ogni elaborazione ulteriore presuppone l'esistenza di dati appartenenti ad un livello più basso.

Si potrebbe dunque partire da un modello bidimensionale come quello rappresentato nella Figura 2.

Comunque, non è facile decidere se un dato particolare sia primario, secondario ecc. Possiamo chiederci, per esempio, se un enunciato fissato nella trascrizione di una conversazione orale sia primario oppure già secondario rispetto alla sua registrazione sul nastro, e *a fortiori* rispetto alla sua manifestazione fonica *in actu* che è stata registrata. Nello stesso modo, una frase nell'edizione di un testo medievale potrebbe considerarsi secondaria rispetto alla frase corrispondente nel manoscritto originale; ecc.

Detto questo, ci sembra più prudente rimaneggiare lo schema riprodotto

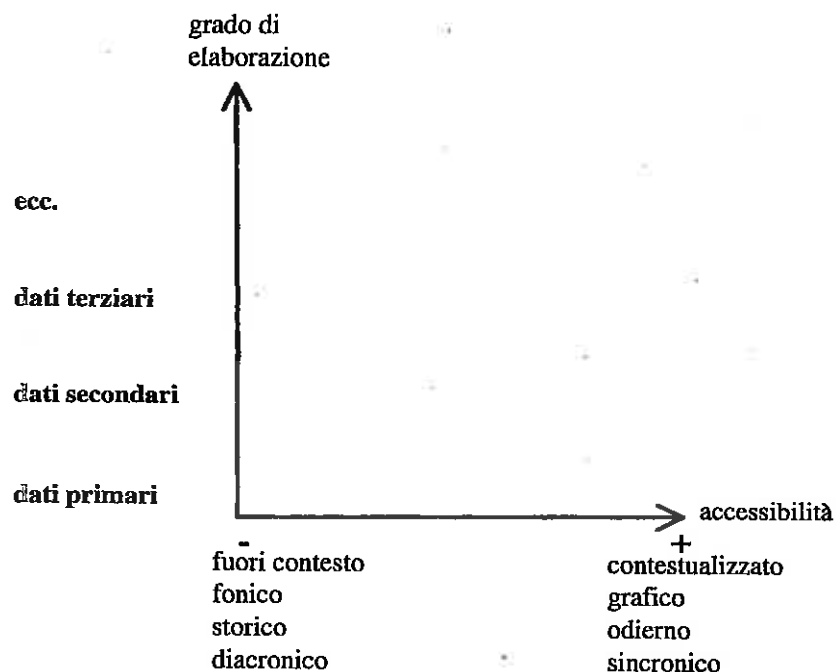


FIGURA 2

in fig. 2, scegliendo come perno teorico un livello zero che corrisponde ad un grado di elaborazione linguistica minima (cfr. fig. 3). Questo livello zero equivale ad una categorizzazione linguistica minima (fonetico-fonologica, grammaticale o lessicale) oppure ad una prima interpretazione semantica, pragmatica ecc. Per giungere al livello zero, c'è eventualmente bisogno di più di una tappa di elaborazione precedente: queste tappe, notate in fig. 3 come  $-1$ ,  $-2$ , ...  $-n$ , possono essere affidate anche a discipline diverse dalla linguistica propriamente detta (paleografia, scienza editoriale, critica testuale, sociologia empirica, fonetica acustica, elettroacustica ecc.). In altre parole, il concetto di dato primario si diluisce completamente a favore di una concezione flessibile per quanto riguarda il numero  $-n$  delle tappe precedenti al livello zero. L'unica cosa che conta è il fatto che il livello zero comporta una categorizzazione o un'interpretazione linguistica minima.

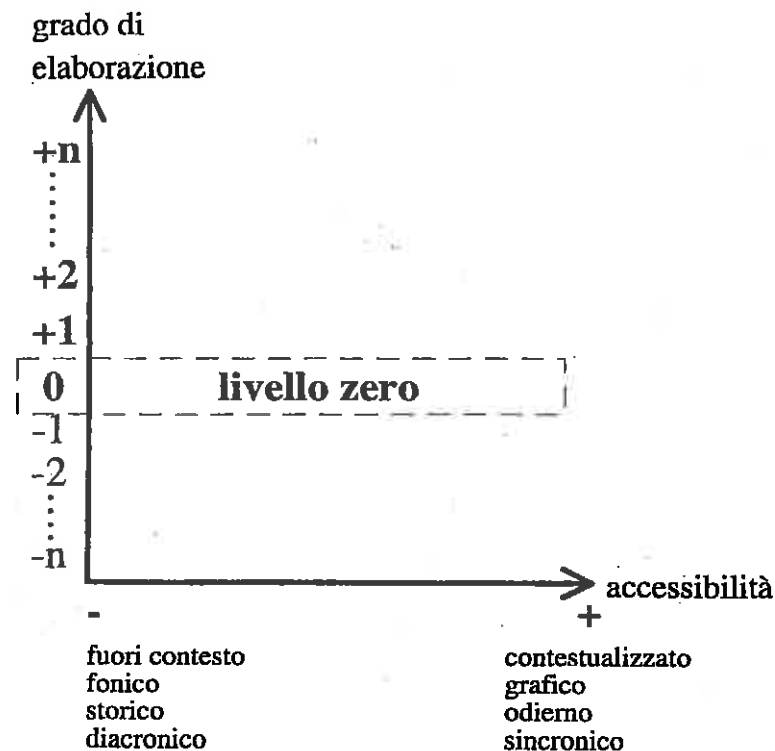


FIGURA 3

Facciamo qualche esempio estratto dal *Dizionario etimologico della lingua italiana* di Cortelazzo - Zolli (1979-88). Vediamo il lemma *gota*:

- (1) **gòta**, s.f. lett. 'guancia' (av. 1292, B. Giamboni; accr. *gotone*: av. 1712, L. Magalotti).  
 • Più che dal lat. *gābata(m)* 'ciotola' (trascritto anche *gauata* e di orig. straniera; e *gote* appaiono col senso proprio di 'scodelle' nella canzone marchigiana di Castra, fiorentino, av. 1250) usato metaforicamente, dal gall. \**gauta* (ID XVIII (1942) 37), accettato quando *bücca(m)*, che, oltre a 'bocca', significava principalmente 'guancia', si specializzò nel solo senso di 'bocca'. *Gota* «è celtico, secondo il Jud (*AStNSpr.*, 124, p. 400), il Dauzat (*Rom.*, 45, 1918-9 p. 253) e il Meyer-Lübke (*Wörter und Sachen*, 12, 1929 p. 8; REW), e in ogni modo irradia certo dalla Gallia: esso lancia punte nella Ladinia occid., Romagna e Toscana occidentale» (Bonfante). [...] (op. cit.: s.v. *gota*)

Troviamo, per prima cosa, delle indicazioni sincroniche odierne del livello zero, riguardanti le categorie morfologiche della parola *gota* («s. f.»), la sua diffusione diatopica («lancia punte nella Ladinia occid...»), il suo valore diafasico odierno («lett.») e il suo significato («'guancia'»). Nella seconda parte, propriamente etimologico-storica, dell'articolo vengono menzionate due filiazioni diacroniche alternative (< lat. *gābata(m)* oppure < gall. \**gauta*). I processi diacronici – e questo esempio ce lo dimostra in modo particolarmente evidente – non sono mai direttamente osservabili in quanto tali. Ogni affermazione prettamente diacronica, traducibile in una formula del tipo

$$x > y$$

si riferisce a due dati *x* e *y*, tra cui viene interpolata una filiazione diacronica (notata «>»), che è necessariamente ipotetica e, in quanto interpretazione dei fatti osservati, corrisponde già al livello zero. Per i «dati» diacronici, i livelli –*n* non esistono!

Nel caso dei dati sincronici (anche storici), la situazione è diversa, perché si può, in linea di principio, risalire ad un livello –*n*, per esempio, ad un'attestazione della parola all'interno di un testo. Nell'articolo (1) gli autori accennano solo indirettamente ad un livello –1, rinviando, tramite l'indicazione di una data e di un autore («av. 1292, B. Giamboni»), alla prima attestazione della parola. In altri articoli dello stesso dizionario, tali dati del livello –1 vengono addirittura citati, come, nel seguente esempio, «Di Cipri si traggono per Damiata mele in quantità e melazzi di zucchero assai»:

- (2) *melàssa*, s. f. 'liquido denso e bruno che rimane dopo che dal succo della canna da zucchero o della barbabietola è stato estratto lo zucchero' (1780, D'Alb.), *melàsso*, s. m. raro 'melassa' (melazzo: 1567, Lud. Guicciardini).
- Fr. *mélasse* (secondo il Batt. documentato dal 1664; in realtà molto precedente: in provz. ant. *melessa* nel 1467, in fr. *meslache* dal 1508), dal lat. *mellāceū(m)* 'sciroppo di miele (mel)', 'mosto cotto, sapa'. La cauta ipotesi di B. Migliorini *Studi Schiaff.* II 713, che il n. sia sorto in Levante e di lì diffuso in Occidente trova un appoggio nelle dichiarazioni dei manuali di mercatura: «Di Cipri si traggono per Damiata mele in quantità e melazzi di zucchero assai» (1503, Di Pasi, cit. in TB). Ma occorrerà indagare meglio sulla parte avuta da Venezia [...]
- (op. cit.: s.v. *melassa*)

Per quanto riguarda i livelli +*n*, bisogna ammettere che i dizionari etimologici tradizionali vi fanno raramente riferimento. Nella seconda parte dell'esempio (1) appare almeno il termine «metaforicamente» che categorizza, al livello +1, un mutamento semantico interpolato al livello zero tra il significato 'ciotola' di un etimo latino ipotetico *gābata(m)* e il significato 'guancia' dell'it. *gota*.

Ancora più raramente i dizionari etimologici tradizionali si avventurano al livello +2. Eccone un esempio:

- (3) *chiàssso*, s. m. 'vicolo corto e stretto' (*classus* nel lat. mediev. di Lucca del 769: Migl. *St. lin.* 72; it. *chiasso*: 1273, NTF), 'postribulo' (av. 1484, L. Pulci), 'rumore, strepito, fracasso' (1572, F. Sasseti). [...]
- L'etim. della parola non è stata del tutto chiarita, anche perché non è del tutto chiaro il rapporto fra i tre sign. ('vicolo', 'postribulo', 'rumore') della medesima. [...] Da parte nostra riteniamo che i tre significati vadano uniti sotto un unico lemma e che l'evoluzione semantica corrisponda perfettamente alla successione cronologica con cui appaiono i tre significati: dal significato di 'vicolo' si è passati a quello di 'postribulo' e da questo a quello di 'rumore'. Quest'ultima evoluzione è confermata dall'analoga evoluzione di *bordello* e *casino*; [...]
- (op. cit.: s.v. *chiasso*)

Al livello zero, gli autori ricostruiscono, per l'it. *chiasso*, un mutamento semantico «dal significato di 'vicolo' [...] a quello di 'postribulo' e da questo a quello di 'rumore'». Analizzando la seconda tappa del mutamento ('postribulo' → 'rumore'), gli autori non esplicitano il livello +1 (dove si dovrebbe parlare senz'altro di 'metonimia'), ma «saltano» al livello +2, riconducendo il mutamento in questione ad un principio in un certo senso cognitivo sottostante a tutta una serie di mutamenti semantici che si osservano anche nelle parole it. *bordello* e *casino*.

#### 4. DATI E RICOSTRUZIONE DIACRONICA

Come dimostrano questi esempi, i «dati» da cui parte il nostro progetto appartengono per la maggior parte al livello zero. I livelli +*n* vengono considerati nei dizionari etimologici tradizionali solo in modo sporadico e poco sistematico (cfr. Blank - Koch 2000, Blank - Koch - Gévaudan 2000). D'altra parte, le informazioni riguardanti i livelli –*n* (attestazioni nei testi ecc.) sono necessariamente sincroniche, mentre l'interpolazione di filiazioni diacroniche che costituisce la base del nostro lavoro corrisponde *per definitionem* al livello zero. Non ci siamo proposti di stabilire nuove etimologie (il che ci costringerebbe magari a scendere ai livelli –*n*), ma di classificare le filiazioni etimologiche ricostruite da altri (livello zero) sulla base di una sistematica cognitiva dei mutamenti lessicali (livello +1) e di ricondurre i mutamenti riscontrati eventualmente a degli schemi cognitivi sottiacenti (livello +2).

Nella prospettiva onomasiologica che è al centro del nostro approccio, constatiamo, ad esempio, che il concetto GUANCIA viene espresso con *mejilla* oppure con *carrillo* in spagnolo moderno, con *falcà* in rumeno e con *kantérdzulkantrèžu* in sardo (cfr. (4)). Lo sp. *mejilla* risale al lat. *maxilla* 'mascella', lo sp. *carrillo* allo sp. ant. *carrillo* 'mascella', il rum. *falcà* al

rum. *falcă* 'mascella' ed il sard. *kantérdzulkantrežu* al lat. volg. *cant(h)erius* 'mascella' (livello zero). Il passaggio GUANCIA ← MASCELLA costituisce un mutamento metonimico (livello +1)<sup>2</sup>.

- (4a) sp. mod. *mejilla* 'guancia' < METON < sp. ant. *mejilla* 'mascella' < lat. *maxilla* 'mascella'
- (4b) sp. *carrillo* 'guancia' < METON < sp. ant. *carrillo* 'mascella' < META < sp. ant. *carrillo* 'carro' ← sp. ant. *carro* < lat. *carrus* 'carro'
- (4c) rum. *falcă* 'guancia' < METON < rum. *falcă* 'mascella' < META/lat. volg. \**falca* < lat. class. *falx* 'falce'
- (4d) sard. *kantérdzu*, *kantrežu* 'guancia' < METON < lat. volg. *cant(h)erius* 'mascella' < META < lat. *cant(h)erius* 'corrente del tetto'
- (dati forniti dal progetto DECOLAR; cfr. Blank - Koch - Gévaudan 2000; cfr. anche Meyer-Lübke 1992: No. 5443, 1721, 3175, 1615)

Riguardo al mutamento metonimico GUANCIA ← MASCELLA, condiviso da (4a), (4b), (4c) e (4d), osserviamo un parallelismo tanto più notevole se consideriamo che il materiale etimologico-lessicale coinvolto è completamente diverso (lat. *maxilla*/vs./lat. *carrus*/vs./lat. *falx*/vs./lat. *cant(h)erius*) (cfr. tab. 3).

TABELLA 3

concetto d'arrivo	dato odierno	meccanismo cognitivo	dato storico / formazione delle parole	concetto di partenza
GUANCIA	sp. mod. <i>mejilla</i> 'guancia'	contiguità (partonomia)	sp. ant. <i>mejilla</i> 'mascella' < lat. <i>maxilla</i> 'mascella'	MASCELLA
GUANCIA	sp. mod. <i>carrillo</i> 'guancia'	contiguità (partonomia)	sp. ant. <i>carrillo</i> 'mascella'	MASCELLA
GUANCIA	rum. <i>falcă</i> 'guancia'	contiguità (partonomia)	rum. <i>falcă</i> 'mascella'	MASCELLA
GUANCIA	sard. <i>kantérdzu</i> , <i>kantrežu</i> 'guancia'	contiguità (partonomia)	lat. volg. <i>cant(h)erius</i> 'mascella'	MASCELLA

<sup>2</sup> METON = mutamento semantico metonimico; META = mutamento semantico metaforico. Per quanto riguarda la metonimia, v. nota seguente.

Potremmo supporre che si tratti di processi poligenetici dovuti a degli schemi cognitivi fondamentali nella percezione della faccia umana: effetti PARTE - TUTTO (partonomia) e contiguità delle parti del corpo, effetti di salienza ecc. (livello +2). Ma visto che abbiamo a che fare con lingue imparentate e con culture vicine l'una all'altra, non è totalmente da escludere che gli schemi cognitivi in questione si siano trasmessi tramite un contatto linguistico e/o culturale. Perciò il nostro progetto va basato su un campione di lingue più ampio e più diversificato. Ampliando, però, il campione, dobbiamo affrontare nuovi problemi. A seconda della qualità della documentazione esistente (livelli -n), la descrizione etimologica di una data lingua o famiglia di lingue (al livello zero) può variare considerevolmente. Malkiel (1994) distingue tre tipi di dizionari etimologici:

- dizionari che si basano su corpora,
- dizionari che si basano su informanti,
- dizionari delle lingue ricostruite, ossia ipotetiche.

Per le lingue romanze possiamo avvalerci di dizionari del tipo (a). Ci troviamo, infatti, in una situazione privilegiata: la documentazione scritta è ottima fin dall'epoca del latino classico (livelli -n). Per la maggior parte delle parole, conosciamo esattamente il loro antecedente latino (greco, germanico ecc.), i loro significati ecc. (livello zero). Rispetto a E = lat. *maxilla*, il caso di sp. mod. *mejilla* (4a) e dei suoi cognati neolatini it. *mascella*, cat. *maixela*, friul. *massèle* e lad. *massèdla* si può rappresentare più o meno così come in fig. 4.

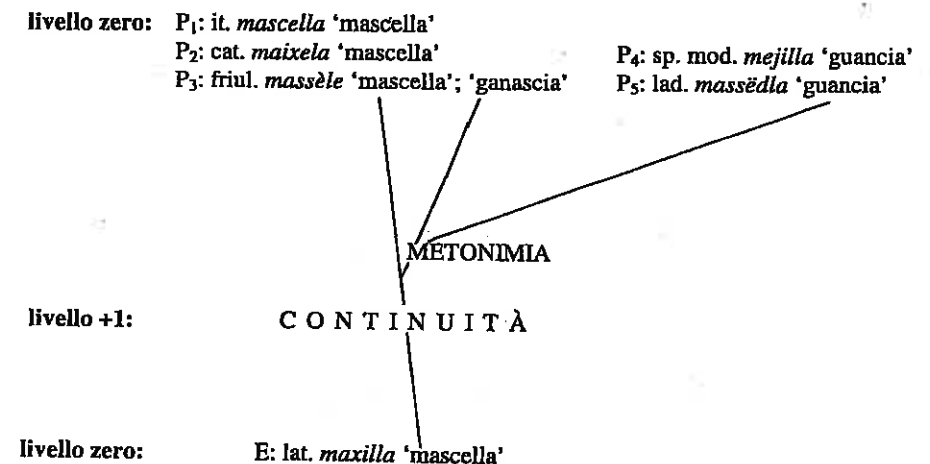


FIGURA 4

Al tipo (h) dei dizionari etimologici appartengono soprattutto i dizionari delle lingue non-europee che fino al Novecento non avevano nessuna tradizione scritta. Un esempio di questo tipo è il dizionario comparativo dell'eschimese (cfr. Fortescue *et alii* 1994) che si basa esclusivamente sugli odierni dati sincronico-comparativi forniti dagli informanti. Ne possiamo ricavare, ad esempio, quanto elencato in (5):

- (5) P<sub>1</sub>: Yupik dell'Alaska centrale *iXcaquq* 'cuore'  
 P<sub>2</sub>: Qawwaraq *iqsaruq* 'pericardio'  
 P<sub>3</sub>: Inuit canadiano orientale *itsayuk* 'epigastrio' (cfr. op. cit.: 144)

Questo materiale ha un carattere completamente diverso da quello presentato in (4) e in fig.4. Le uniche informazioni disponibili, al livello zero, sono le parole P<sub>1</sub>-P<sub>3</sub> delle tre lingue eschimesi moderne indicate. Solo al livello +1 constatiamo delle somiglianze foniche tra le parole P<sub>1</sub>-P<sub>3</sub> e delle relazioni semantiche (di contiguità) tra i concetti che esse designano, fatto, questo, che deve avere portato gli autori a raggruppare P<sub>1</sub>-P<sub>3</sub> (assieme ad altre parole) sotto lo stesso lemma. Sembra abbastanza probabile, infatti, che P<sub>1</sub>, P<sub>2</sub> e P<sub>3</sub> siano connesse tra di loro da un legame di parentela. Questo stato di fatto si potrebbe rappresentare così come in fig. 5.

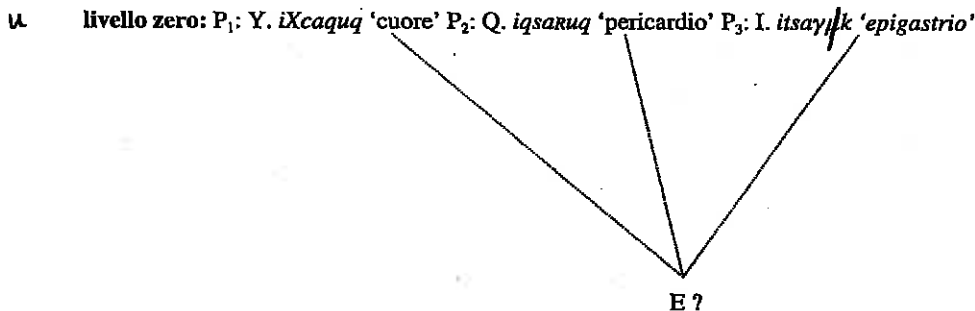


FIGURA 5

Gli autori, però, non si fermano qua e raggruppano P<sub>1</sub>-P<sub>3</sub> (e altre parole) sotto un lemma comune:

- (6) E: proto-eschimese *ircaqur* 'parte intorno al cuore' (ibid.)

Ovviamente questo etimo corrisponde al punto di convergenza diacronica E in fig.5, cioè ad una forma ricostruita (che dovrebbe essere accompagnata da un asterisco: \*). Questa ricostruzione non è un «dato» appartenente al livello zero (e ancora meno ad un livello -n), bensì il frutto di ragionamen-

ti che si svolgono al livello +1. Ammesso che la ricostruzione fonica sia metodologicamente salda, bisogna comunque segnalare che la ricostruzione di un significato «di base» è impresa estremamente problematica. Nel caso presente gli autori ci offrono un significato di «convergenza» ('parte intorno al cuore') che ricopre in un certo senso tutti i significati accessibili alla superficie sincronica odierna: siccome essi sono legati tra di loro tramite relazioni di contiguità concettuale, si è scelto un significato di convergenza che circonda più o meno esattamente il *frame*<sup>3</sup> cognitivo che unisce tutti i concetti espressi dalle odierne parole P<sub>1</sub>-P<sub>3</sub>. Dal punto di vista empirico un tale significato di convergenza è vano, perché non è un dato equiparabile ai significati di P<sub>1</sub>-P<sub>3</sub> (livello zero), ma dipende interamente da essi (livello +1). Per il nostro lavoro questa osservazione significa che in tali casi non possiamo partire da una filiazione diacronica E > P<sub>1</sub>-P<sub>3</sub> al livello zero per poi distillarne, al livello +1, delle relazioni semantiche (p.es. metonimiche), ma che, almeno dal punto di vista semantico, dobbiamo fare a meno dell'entità ipotetica E. Ciò nonostante, il materiale puramente sincronico raccolto in (5) merita il nostro interesse perché abbiamo il diritto di prendere atto, al livello +1, delle relazioni di contiguità che esistono tra i concetti designati dalle parole P<sub>1</sub>-P<sub>3</sub> e di interpretarle in chiave diacronica (senza nessun ricorso all'entità ricostruita E: cfr. fig. 6).

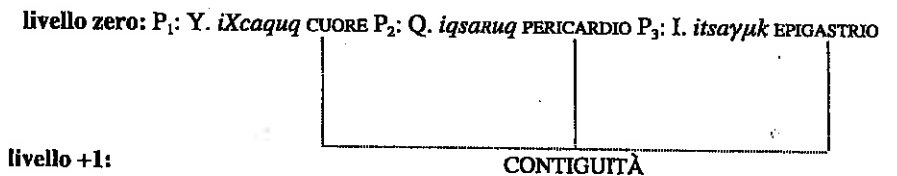


FIGURA 6

Possiamo poi, all'occorrenza, integrare i risultati di questa analisi del livello +1 nel riconoscimento di schemi cognitivi più astratti (al livelli +2) che spiegano eventualmente dei mutamenti lessicali ricorrenti e poligenetici.

Queste riflessioni s'inquadrano perfettamente nella concezione più astratta che si delinea nei lavori legati ai nostri progetti di ricerca nell'ambito della semantica cognitiva diacronica, nonché sincronica. Invece di parlare, per esempio, sempre di 'metonimia', è preferibile in certi casi identificare un complesso di mutamenti lessicali (oppure di rapporti semantici sincronici)

<sup>3</sup> Per quanto riguarda il rapporto tra la metonimia, i *frame* cognitivi e le relazioni di contiguità, cfr. Blank (1997: 230 ss.); Koch (1999b).



«coinvolgendo una relazione di contiguità concettuale» (e in modo simile per la metafora rispetto alla similarità metaforica, ecc.).

Prima di concludere bisogna brevemente accennare ai dizionari etimologici appartenenti al tipo (c), i quali si basano su dati misti; analizzano, cioè, dei corpora, ma considerano anche dei dati sincronico-comparativi. Si pensi, per esempio, alla ricostruzione dell'indoeuropeo o del proto-austronese (Pokorny 1959/69; Tryon 1995). Colpisce il fatto che in questo ambito la situazione dei «dati» spesso non sia molto diversa da quella descritta a proposito di (5) e (6), come lo dimostra il seguente esempio tratto dal dizionario indoeuropeo più rinomato:

- (7) P<sub>1</sub>: lat. *botulus* 'intestino'  
 P<sub>2</sub>: got. *qībus* 'stomaco; grembo materno'  
 P<sub>3</sub>: isl. ant. *kviðr* 'ventre; grembo materno'  
 P<sub>4</sub>: a.ted. ant. *qūiti* 'vulva'

E: ie. *gʷet-* 'gonfiore, forma arrondata' (cfr. Pokorny 1959/69: 481)

Anche in questo caso, dobbiamo trascurare, dal punto di vista semantico, l'entità ipotetica E. Notiamo, però, al livello +1 le relazioni di contiguità (e di similarità) che esistono tra i concetti designati dalle parole P<sub>1</sub>-P<sub>4</sub> e le interpretiamo in chiave diacronica, identificando un complesso di mutamenti lessicali coinvolgenti una relazione di contiguità oppure di similarità concettuale.

Ciò che ci sembra importante sottolineare è il fatto che l'indogermanistica, scienza pilota delle ricerche etimologiche e comparative, spesso non si distacca molto dalla ricerca etimologica nell'ambito di lingue abbastanza «esotiche»<sup>4</sup>.

## 5. CONCLUSIONE

Ciò che abbiamo voluto evidenziare è che il carattere particolare dei dati indagati nel nostro progetto di ricerca ci costringe ad una riflessione approfondita sulla raccolta dei dati e sulla loro tipologia. Essendo i nostri dati di natura lessicale, il campionamento sarà regolato da criteri differenti da quelli utilizzati nell'ambito grammaticale. Per quanto riguarda la tipologia dei dati, abbiamo sviluppato un modello bidimensionale basato sui due assi dell'accessibilità e del grado di elaborazione. Abbiamo situato il nostro lavoro rispetto ad un livello zero, dimostrando in cosa consiste l'elaborazione di dati

<sup>4</sup> Questo vale, in fin dei conti, anche per certe etimologie slave ricostruite, come dimostra l'esempio protoslavo \**usta* nella tab. 1.

lessicali (diacronici) del livello +1 (interpretazione dei mutamenti in termini di 'metonimia', 'metafora' ecc.) e del livello +2 (riconoscimento di schemi cognitivi soggiacenti). Finalmente ci siamo resi conto del fatto che i significati ricostruiti non appartengono al livello zero e pertanto risultano essere empiricamente vani. Questo, però, non ci impedisce di stabilire, in termini più astratti ('contiguità', 'similarità metaforica' ecc.), legami diacronici tra esiti sincronici in lingue imparentate tra di loro, senza ricorrere a dei significati di «convergenza» ricostruiti e artificiali.

Tutto sommato, si può pretendere senza esagerazione che una casistica di dati particolarmente intricata incoraggi anche la teorizzazione di una problematica dei dati linguistici in genere.

## BIBLIOGRAFIA

### Dizionari

Buck, C. D. (1949), *A Dictionary of Selected Synonyms in the Principal Indo-European Languages: A Contribution to the History of Ideas*, Chicago, University Press.

Cortelazzo, M. - Zolli, P. (1979-88), *Dizionario etimologico della lingua italiana*, Bologna, Zanichelli.

Fortescue, M. - Jacobson, S. - Kaplan, L. (1994), *Comparative Eskimo Dictionary. With Aleut Cognates*, Fairbanks, Alaska Native Language Center.

Machek, V. (1957), *Etymologický Slovník Jazyka Českého a Slovenského*, Praha, Československé Akademie Věd.

Meyer-Lübke, W. (1992), *Romanisches etymologisches Wörterbuch*, Heidelberg, Winter.

Pokorny, J. (1959/69), *Indogermanisches etymologisches Wörterbuch*, 2 vol., Bern, Francke.

Schröpfer, J. (1979 ss.) (a cura di), *Wörterbuch der vergleichenden Bezeichnungen*, Heidelberg, Winter.

Trubačev, O. N. (1979) (a cura di), *Ėtimologičeskij slovar' slavjanskich jazykov. Praslavjanskij leksičeskij fond*, fasc. 6, Moskva, Nauka.

Tryon, D. T. (1995) (a cura di), *Comparative Austronesian Dictionary: An Introduction to Austronesian Studies*, Berlin-New York, Mouton de Gruyter.

Vasmer, M. (1953), *Russisches etymologisches Wörterbuch*, Heidelberg, Winter.

Vernay, H. (1991 ss.), *Dictionnaire onomasiologique des langues romanes*, Tübingen, Niemeyer.

### Bibliografia dei titoli citati

Andersen, E. S. (1978), *Lexical universals of body-part terminology*, in Greenberg, J.

- 733 H. (a cura di), *Universals of Human Language*, Vol. III, Stanford, University Press: 295-368.
- Berincà, P. (1994), *Linguistica e dialettologia italiana*, in Lepschy, G. C. (a cura di), *Storia della linguistica*, vol. III, Bologna, Il Mulino: 524-644.
- Blank, A. (1997), *Prinzipien des lexikalischen Bedeutungswandels am Beispiel der romanischen Sprachen*, Tübingen, Niemeyer.
- (1998), *Der Kopf in der Romania und anderswo – ein metaphorisches (und metonymisches) Expansions- und Attraktionszentrum*, in Gil, A. - Schmitt, Chr. (a cura di), *Kognitive und kommunikative Dimensionen der Metaphorik in den romanischen Sprachen*, Bonn, Hilten: 11-32.
- Blank, A. - Koch, P. (2000), *La conceptualisation du corps humain et la lexicologie diachronique romane*, in Dupuy-Engelhardt, H. - Montibus, M. J. (a cura di), *La lexicalisation des structures conceptuelles*, Reims, Presser Universitaires de Reims: 43-62.
- Blank, A. - Koch, P. - Gévaudan, P. (2000), *Onomasiologie, sémasiologie et l'etymologie des langues romanes: esquisse d'un projet*, in Englebert, A. et alii (a cura di), *Actes du XXI<sup>e</sup> Congrès International de Linguistique et de Philologie Romanes. IV: Des mots aux dictionnaires*, Tübingen, Viemeyer: 103-104.
- Corrà, L. (1981), *Contributo alla bibliografia onomasiologica. Dominio italiano*, «La ricerca dialettale», 3: 392-478.
- Haspelmath, M. (1997), *Indefinite Pronouns*, Oxford, Clarendon.
- Koch, P. (1997), *La diacronia quale campo empirico della semantica cognitiva*, in Carapezza, M. - Gambarara, D. - Lo Piparo, F. (a cura di), *Linguaggio e cognizione*, Roma, Bulzoni: 225-246.
- (1999a), *Cognitive aspects of semantic change and polysemy: the semantic space HAVE/BE*, in Blank, A. - Koch, P. (a cura di), *Historical Semantics and Cognition*, Berlin-New York, Mouton de Gruyter: 279-305.
- (1999b), *Frame and contiguity. On the cognitive bases of metonymy and certain types of word formation* in Panther, K. - U. - Radden, G. (a cura di), *Metonymy in Language and Thought*, Amsterdam-Philadelphia, Benjamins: 139-167.
- (1999c), *TREE and FRUIT: a cognitive-onomasiological approach*, «Studi Italiani di Linguistica Teorica/Applicata», 28/2: 331-347.
- (in stampa), *Pour une approche cognitive du changement sémantique lexicale: aspect onomasiologique*, in François, J. (a cura di), *Mémoire 2000 de la Société de Linguistique de Paris*, Leaven, Peeters.
- Malkiel, Y., (1994), *Etymology*, in *The Encyclopedia of Language and Linguistics*, Vol. 3, Oxford ecc., Pergamon: 1168-1178.
- Nichols, J. (1992), *Linguistic Diversity in Space and Time*, Chicago, University of Chicago Press.
- Oehl, W. (1922), *Elementare Wortschöpfung*, in *Miscellanea linguistica dedicata a Hugo Schuchardt*, Ginevra, Olschki: 75-115.
- Quadri, B. (1952), *Aufgaben und Methoden der onomasiologischen Forschung*, Bern, Francke.
- Rijkhoff, J. - Bakker, D. - Hengeveld, K. - Kahrel, P. (1993), *A method of language sampling*, «Studies in Language» 17: 169-203.
- Schladt, M. (1997), *Kognitive Prinzipien von Körperteilvokabularien in kenianischen Sprachen*, Köln, Universität Köln.
- Tagliavini, C. (1949), *Di alcune denominazioni della 'pupilla'. Studio di onomasiologia, con speciale riguardo alle lingue camito-semitiche e negro-africane*, in *Scritti minori*, Bologna, Pàtron 1982: 529-568.
- Wilkins, D. P. (1996), *Natural tendencies of semantic change and the search for cognates*, in Durie, M. - Ross, M. (a cura di), *The Comparative Method Reviewed*, Oxford, Oxford University Press: 264-304.