

Silvana Ferreri

L'alfabetizzazione lessicale
Studi di linguistica educativa



Copyright © MMV
ARACNE EDITRICE S.r.l.

www.aracneeditrice.it
info@aracneeditrice.it

00173 Roma
via Raffaele Garofalo, 133 A/B
(06) 93781065

ISBN 88-548-0137-2

*I diritti di traduzione, di memorizzazione elettronica,
di riproduzione e di adattamento anche parziale,
con qualsiasi mezzo, sono riservati per tutti i Paesi.*

I edizione: dicembre 2005

INDICE

| | |
|-------------------------------------------|-----|
| <i>Premessa</i> | 7 |
| TEMI | |
| I. Conoscenze lessicali individuali | 11 |
| II. Strumenti lessicografici | 43 |
| III. Memoria e apprendimento del lessico | 65 |
| PROBLEMI | |
| IV. Parole e parlanti | 83 |
| V. Piano di alfabetizzazione lessicale | 95 |
| VI. Parole e termini | 111 |
| APPLICAZIONI | |
| VII. Parole tra scuola e università | 125 |
| VIII. Definire le parole delle discipline | 139 |
| IX. Sinonimia e variazione | 149 |
| X. Parole per l'italiano L2 | 161 |
| XI. Marche d'uso e dizionari bilingui | 173 |
| <i>Bibliografia</i> | 193 |
| <i>Indice dei nomi</i> | 203 |
| <i>Indice analitico</i> | 205 |

Premessa

Raccoglio in questo volume studi e interventi sul lessico presentati e pubblicati negli ultimi due anni in sedi italiane e straniere¹. Essi si radicano nella consapevolezza teorica sempre più diffusa della centralità che il lessico ha nel funzionamento delle lingue storico-naturali e nei processi di apprendimento non solo linguistico. I lavori sono stati in molti casi riscritti, in altri tagliati, smontati e rimontati per dare maggiore funzionalità al volume nel suo insieme. Tuttavia per conservare coerenza ai singoli testi ho preferito lasciare in taluni casi qualche ripetizione: forse aiuta a intendere l'importanza di alcuni nodi concettuali e fattuali imbattersi in essi e riconsiderarli lungo percorsi diversi.

I fatti di lingua sono qui analizzati in una prospettiva che possa risultare utile a chi insegna lingue e a chi in altri ambiti disciplinari tenga conto, come da gran tempo chiediamo, della dimensione linguistica inerente a ogni insegnamento e apprendimento: è la prospettiva che oggi diciamo *linguistica educativa*, assunta nel sottotitolo del volume.

Dedico questo libro a Tullio De Mauro: per il molto che gli devo il poco che posso.

s.f.

Roma, dicembre 2005

¹ I lavori hanno le seguenti origini: *L'estensione delle conoscenze lessicali individuali*, in Tullio De Mauro – Isabella Chiari (a cura di), *Parole e numeri. Analisi quantitative dei fatti di lingua*, Aracne, Roma 2005: 307-334; *Quantità dei lemmi nei dizionari*, in Tullio De Mauro – Isabella Chiari (a cura di), *Parole e numeri. Analisi quantitative dei fatti di lingua*, Aracne, Roma 2005: 297-306; relazione «Costruire significati per scrivere», svolta nell'ambito del seminario del GISCEL sulla Scrittura, Fiuggi, gennaio 2004; *Relazioni di senso e frequenza delle parole*, in Anna Laura Lepschy e Anna Rita Tamponi (a cura di), *Prospettive sull'italiano come lingua straniera*, Atti del convegno *Italian as a Foreign Language*, London University College, Londra 28-29 novembre 2003, Guerra, Perugia 2005: 47-62; *Fare i conti con le parole*, in Miriam Voghera, Grazia Basile, Anna Rosa Guerriero (a cura di), *E.L.I.C.A. Educazione linguistica e conoscenze per l'accesso*, Guerra, Perugia, 2005, pp. 180-194; *L'esercizio della definizione*, in «Vita scolastica», anno 58, 18 (20 giugno 2004), pp. 23-25; *Scegliere i sinonimi parlando e scrivendo*, in Tullio De Mauro, *Dizionario dei sinonimi e contrari*, Paravia, Torino 2002, pp. 1092-1105; *Parole tra quantità e qualità*, relazione tenuta al XIII Convegno Nazionale Gisel, *Il linguaggio dall'infanzia all'adolescenza: tra italiano, dialetto e italiano L2*, Lecce, 23, 24, 25 aprile 2004, (in corso di stampa) in Tempesta Immacolata e Maggio Maggio (a cura di), *Il linguaggio dall'infanzia all'adolescenza*, Franco Angeli, Milano; *Questioni di metodo o questioni linguistiche? Note per una linguistica educativa*, in AA.VV., *Insegnare italiano in Giappone. Problemi e prospettive. Nihon ni okeru Itariago Kyoitu. Mondai to tenbo*, Atti del 1° seminario di studi, Istituto italiano di cultura, Tokyo 2005, pp. 12-37; *Memoria e apprendimento del lessico*, relazione tenuta al convegno su «Comunicare le memorie: Itinerari della conoscenza», Roma-Viterbo 6-8 novembre 2003, (in corso di stampa) in «Heteroglossia», Università degli studi di Macerata; *Servono le marche d'uso in un dizionario bilingue?* relazione al colloquio su «Progetto di dizionario italiano-svedese». 1° Colloquio, Stoccolma, 10-12 febbraio 2005, (in corso di stampa negli Atti).

Temi

I

Conoscenze lessicali individuali

1. Carattere incrementale del lessico

Nel considerare la natura del lessico¹ e il suo ruolo nel costituirsi della competenza linguistica, teorici del linguaggio, linguisti applicati e glottodidatti spesso esprimono posizioni divergenti in conseguenza dell'orientamento teorico o dell'approccio disciplinare di elezione. Tuttavia, nelle loro analisi si rintraccia un elemento comune relativo ad un carattere del lessico: essi concordano infatti nel riconoscere la natura incrementale della conoscenza lessicale², per altro suscettibile nel tempo anche di forte contrazione (cfr. Jackendoff 2002: 241; Na-

¹ Per lessico si intende l'insieme delle parole effettivamente presenti in una lingua; nell'uso corrente due parole equivalenti sono *dizionario* e *vocabolario*: qui si indicherà con *dizionario* la raccolta e rappresentazione, solitamente alfabetica, di una parte più o meno estesa, ma, come si vedrà, necessariamente limitata del lessico, e con *vocabolario* una porzione effettiva di lessico limitata all'uso di un autore, di una fascia sociale, di un campo ecc. (il *vocabolario di Montale*, il *vocabolario dei giovani*, il *vocabolario della psicanalisi*).

² Per definire la natura della conoscenza lessicale sarebbe più appropriato parlare di carattere variabile, per ricordare che accanto alle possibilità di ampliamento di cui si discute approfonditamente in letteratura, esiste una spinta opposta connessa con il contrarsi e ridursi delle parole a disposizione del parlante. Questo aspetto è poco considerato nelle ricerche lessicali e se ne ha contezza solo indirettamente in lavori collaterali.

tion 1990: 33; Melka 1997: 99; Schmitt, Meara 1997: 20; Schmitt 2000: 4 e 137). Gli studiosi convengono sulla natura processuale dell'acquisizione lessicale e concordano sul fatto che, a differenza della sintassi o della fonologia, il cui possesso si stabilizza entro un periodo di tempo definito, la conoscenza delle parole di qualsiasi lingua, nativa e/o seconda, avvenga per gradi e non possa dirsi mai conclusa per la natura stessa della componente lessicale.

Il lessico soggiace alla mutevolezza delle necessità ideative ed espressive dei membri di una comunità e affida alla indefinitezza delle sue unità e all'estensibilità dei suoi significati il carico della sua indeterminabilità. Trasposti sul piano della quantificazione della competenza lessicale, questi caratteri rendono ardui l'accertamento dell'entità della conoscenza acquisita spontaneamente e la progettazione di un apprendimento lessicale programmato. Sia sotto il profilo squisitamente numerico sia per la pluralità degli aspetti connessi al senso che si attribuisce al conoscere una parola, determinare in quantità e qualità l'entità del bagaglio lessicale di ognuno si presenta come un problema di non facile risoluzione.

Vieppiù complesso appare programmare un ampliamento delle complessive capacità di codifica e decodifica lessicale in ambiente istituzionalizzato, perché sul versante delle parole ogni ipotesi di intervento richiede che dati quantitativi ed elementi qualitativi vengano considerati congiuntamente sia in fase di progettazione di un possibile sviluppo, inteso come arricchimento dell'input per favorire l'acquisizione naturale o come strategie didattiche per incentivare l'apprendimento guidato, sia nelle fasi di valutazione della consistenza della competenza colta in senso longitudinale. Appare infatti difficile mettere a fuoco lo specifico riguardante l'incremento della conoscenza delle parole e del loro uso se non si conviene su che cosa può voler dire includere il lessico in un percorso di apprendimento e non si conviene su possibili stadi di espansione delle conoscenze.

In un saggio del 1997, Laufer lamenta che la ricerca sul vocabolario in quanto tale è trascurata a favore di lavori collaterali concernenti la gestione dell'apprendimento lessicale. Egli imputa tale carenza alla natura del lessico che per il suo carattere asistemico non si presta al tipo di codificazione formale che si ottiene nei sistemi chiusi come, a suo avviso, in quello fonologico o sintattico. Pur tuttavia i lavori

sull'apprendimento lessicale degli ultimi anni sono riusciti a evidenziare che a fronte del carattere aperto dell'insieme lessicale è possibile riscontrare regolarità che permettono di ipotizzare una gradualità della conoscenza sia quantitativa sia qualitativa delle parole.

Sullo sfondo di una ricognizione dei dati sulla consistenza del vocabolario individuale, nel lavoro si sostiene la tesi della correlazione tra ampiezza quantitativa e profondità qualitativa, come fattore che influenza favorevolmente l'espansione della conoscenza lessicale. Seguendo i principi e le indicazioni della *Educazione Linguistica* sul ruolo delle parole nello sviluppo del linguaggio (cfr. De Mauro 1980) e le implicazioni del *Lexical Approach* (cfr. Lewis 1993) al lessico viene riconosciuta una centralità nel costituirsi della competenza linguistica. Nel corso del lavoro si discuterà la possibilità di prefigurare piani di incremento lessicale quali-quantitativi connessi con le necessità dell'elaborazione semantica e della conservazione in memoria delle conoscenze lessicali.

2. Primi stadi di sviluppo lessicale

2.1. Circoscrivere l'universo lessicale

L'estensione del vocabolario individuale è stata misurata nel corso del tempo in varie lingue con lavori di diverso spessore e consistenza, che tuttavia non riescono a fornire un quadro complessivo dei risultati e non offrono elementi per valutare lo sviluppo della competenza lessicale nel tempo. Le singole ricerche quantitative sono anche di non facile reperibilità, talora datate, difficilmente comparabili sul piano metodologico e in qualche caso persino contraddittorie. La loro problematicità scaturisce dall'oggetto stesso dell'accertamento: si misura il numero di parole conosciute senza precisare il valore che si attribuisce a 'parola' (vedi in questo capitolo, par. 4.1) e senza specificare l'ordine di grandezza con cui lo si confronta. Per ovviare a queste difficoltà appare necessario chiarire i sensi che si attribuiscono ai termini in uso: parola, conoscenza delle parole, competenza lessicale, universo lessicale di riferimento. Si prenda ad esempio il lessico di una qualsiasi lingua. Per farsene una idea bisogna ridurlo, limitarlo, chiuderlo e, così facendo, lo si stravolge perché esso è per sua natura vasto, illi-

mitato, aperto, indeterminabile. Tuttavia, per rappresentarlo, il lessico richiede che si circoscriva il suo universo in modo da evidenziarne anche le forze agenti al suo interno. Analoga operazione di delimitazione va condotta se si vuole procedere a definire tanto la consistenza quantitativa delle parole della lingua nativa che gruppi omogenei di parlanti acquisiscono spontaneamente quanto anche quantità programmate di parole, porzioni di lessico da sviluppare progressivamente in uno con l'apprendimento delle materie in rapporto all'età e agli ordini di scuola. Per questa duplice necessità di chiusura degli insiemi lessicali è utile assumere come riferimento i dizionari, da quelli più ampi e generali (vedi cap. II) a quelli più contenuti nel numero di lemmi, progettati per scopi pedagogici.

2.2 Parole in quantità

Il vocabolario individuale è oggetto di studio in più lingue e in più campi di ricerca. Lunghi dall'essere un elemento di ricchezza investigativa e interpretativa, la pluralità linguistica e di approcci e la separazione in cui operano i diversi campi disciplinari finiscono per rendere difficile l'accesso a molte informazioni preziose per valutare l'evolversi dell'ampiezza lessicale.

I primi dati sull'acquisizione di parole provengono prevalentemente dagli studi di psicologia evolutiva e sono orientati ad esaminare l'evolversi della comparsa di nuove parole nei bambini dai primi mesi di vita fino ai sei anni badando soprattutto alla produzione. Le prime fasi di acquisizione delle parole della lingua nativa operando con numeri ridotti di vocaboli permettono di cogliere tendenze dello sviluppo e stadi evolutivi in grado di favorire la riflessione anche sui livelli più avanzati di età.

Le tabelle sullo sviluppo del vocabolario infantile inglese in produzione evidenziano una correlazione tra età e ampiezza del vocabolario: a parità di età i bambini mostrano di possedere in media un bagaglio lessicale di pari grandezza. Nella media si riscontra un incremento di parole correlato con l'età ma la variazione tra i singoli individui di una stessa fascia di età è tanto ampia da far ritenere preferibile accomunare i bambini in base al numero di parole conosciute più che al numero di mesi compiuti. Al di là di differenze individuali però, l'acquisizione

delle parole in produzione si manifesta con una progressione regolare, generalizzabile, che passa dalla prima parola prodotta intorno al primo anno di età fino ad un consistente numero di vocaboli adoperati a sei anni.

Tabella 1 - Età e incremento lessicale produttivo - lingua inglese

| <i>Età</i> | <i>Parole</i> |
|------------|---------------|
| 10 mesi | Prima parola |
| 12 mesi | 3 parole |
| 18 mesi | 20 parole |
| 2 anni | 250 parole |
| 3 anni | 900 parole |
| 4 anni | 1.500 parole |
| 5 anni | 2.000 parole |
| 6 anni | 2.500 parole |

Il valore relativo della media sulla cui base si valutano gli scarti individuali si apprezza se si mettono a confronto due studi diversi per quantità e metodologia, centrati su anglofoni nativi, reperibili al sito <http://www.ling.upenn.edu>. In essi si evidenzia la velocità dello sviluppo lessicale colto su bambini anglofoni, ma si coglie il *range* di variazione sia per quanto riguarda l'età sia per il numero di parole. Nelson 1973 e Fenson 1993³ variano nell'entità del campione e nella metodologia di accertamento: il primo studio è frutto di registrazioni giornaliere effettuate direttamente dalle mamme su tutta la produzione verbale dei loro bambini; il secondo è l'esito del riconoscimento condotto su una lista di vocaboli rispetto a cui le madri indicano quali sono le parole note ai loro figli. Ciò che si evidenzia con nettezza è il costituirsi della media: secondo Nelson la quantità di 10 parole si raggiunge a 15 mesi con una fascia che va da 13 a 19 mesi; in Fenson la stessa quantità è indicata a 13 mesi, ma con una variazione tra gli 8 e i

³ Per i riferimenti ai lavori riportati, si consulti il sito www.ling.upenn.edu.

16 mesi. Analogo divario si ha nella quantità di parole registrate a due anni.

Tabella 2 - Confronto tra ricerche

| <i>Tappe</i> | <i>Nelson 1973 (18 bambini)</i> | <i>Fenson 1993 (1.789 bambini)</i> |
|--------------------------|-------------------------------------|----------------------------------------|
| 10 parole | 15 mesi (fascia 13-19) | 13 mesi (fascia 8-16) |
| 50 parole | 20 mesi (fascia 14-24) | 17 mesi (fascia 10-24) |
| Vocabolario a 24 mesi | 186 parole (fascia 28-436) | 310 parole (fascia 41-668) |

Fonte: www.ling.upenn.edu

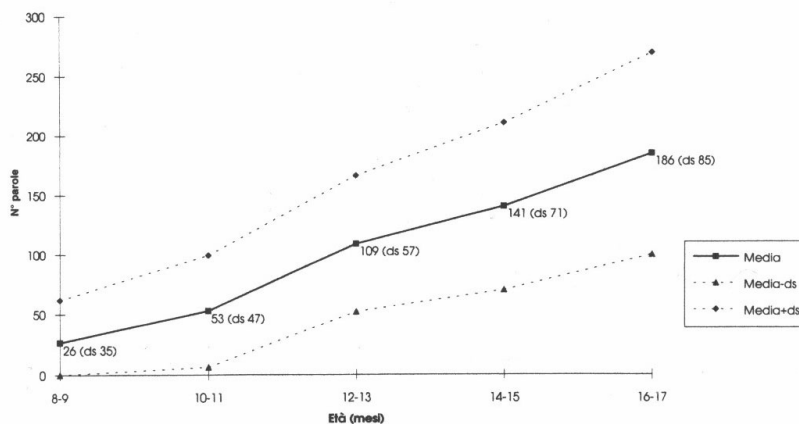
In conseguenza della metodologia impiegata, i due studi divergono per entità del campione e quantità di vocaboli, ma in entrambi si evidenzia l'ampia banda di oscillazione tra i dati rilevati. La disparità tra dati minimi e massimi registrati emerge dall'indicazione della fascia entro i cui limiti si ricava la media: 186 parole sono il punto di equilibrio tra gli estremi di 28 e 436 parole (Nelson); la media di 310 opera entro confini che vanno da 41 parole a 668 (Fenson). L'incidenza della variazione individuale si coglie con nettezza, ma con altrettanta evidenza risalta l'incremento lessicale che si determina a due anni, che sebbene non venga più considerata da molti come la fase dell'esplosione lessicale è certamente il periodo in cui si registra una crescita esponenziale di vocaboli nuovi.

Per la produzione di parole italiane, Caselli e Volterra (1999) riportano dati leggermente superiori a quelli indicati nella tabella per l'inglese: a 12 mesi i bambini producono in media 8 parole, a 16 mesi 32 parole, a 17-18 mesi 54 parole, a 19-21 mesi circa 130, a 2 anni e mezzo circa 400 parole, con un'ampia banda di oscillazione. La grande variabilità individuale, sottolineata da tutti i ricercatori, si apprezza maggiormente nella produzione di parole a 20-21 mesi in cui, a fronte di una media di 130 parole, vi sono bambini che producono un numero minimo di parole pari a 7 e altri che si destreggiano con un vocabolario di 484 parole (cfr. Caselli e Casadio, 1995: 76). I dati numerici

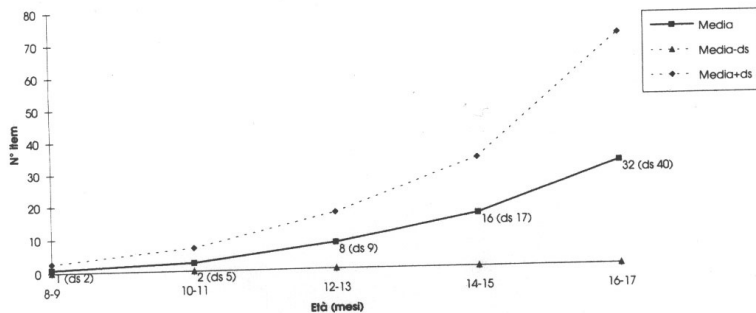
medi, dunque, vanno considerati come linee di tendenza nello sviluppo degli individui più che come stadi di un processo obbligato su cui misurare gli scarti individuali. Essi tuttavia rappresentano un buon indicatore per cogliere precocemente eventuali scostamenti patologici.

Ma dagli studi di Caselli e Volterra emerge anche la diversificazione quantitativa tra la comprensione, da altri trascurata, e la produzione. La ricezione, ben lontana dall'essere attività passiva, precede e prepara la più immediatamente percettibile capacità produttiva, definita attiva per la tangibilità delle sue manifestazioni. Le differenze quantitative, nettamente a favore della comprensione, consentono di cogliere l'importanza della fase di ascolto e ricezione nel costituirsi della competenza lessicale. L'incidenza della ricezione sulla produzione agisce con un differenziale che opera tanto a favore del valore numerico delle parole – si conoscono più parole in comprensione di quante non se ne usino produttivamente – quanto in termini di età: un livello pari per estensione di vocaboli si raggiunge in produzione con uno scarto di circa 12 mesi.

Grafico 1 - Comprensione di parole da 8-9 mesi a 16-17

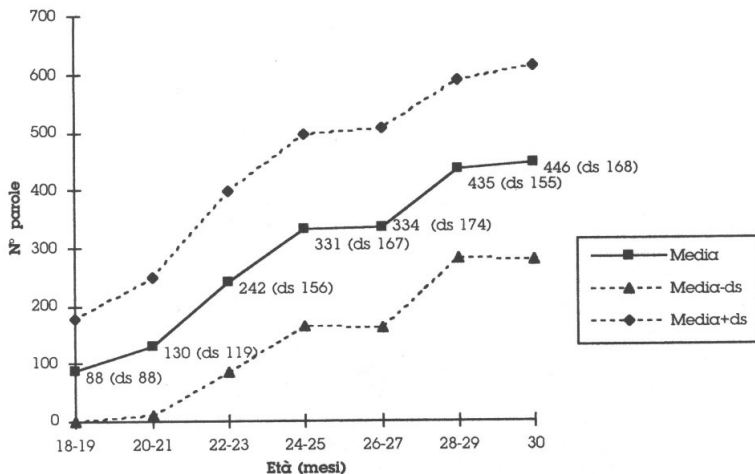


Fonte: Caselli e Casadio (1995: 130)

Grafico 2 - Produzione di parole da 8-9 mesi a 16-17

Fonte: Caselli e Casadio (1995: 133)

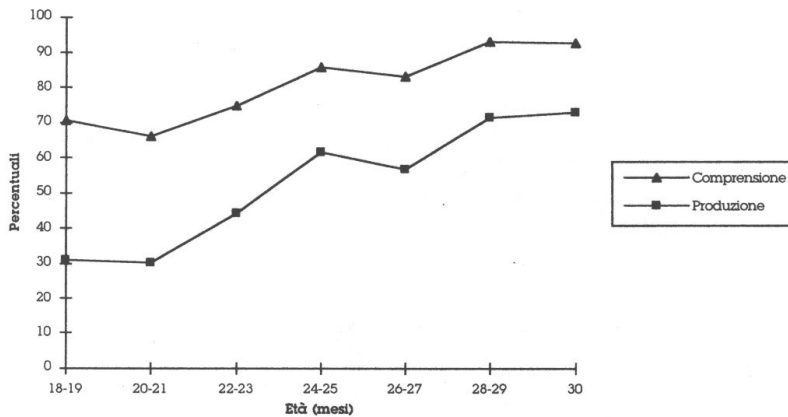
Nel passaggio da un anno e mezzo a 2 anni e mezzo (30 mesi) si coglie la forte crescita in produzione del numero delle parole. L'andamento della fase produttiva (v. oltre Grafico 3) richiama quello della ricezione della fascia di età immediatamente precedente (cfr. Grafico 1), che indica come sia necessario cambiare fascia di età per apprezzare il lavoro di stabilizzazione delle prime parole e il loro uso attivo.

Grafico 3- Produzione di parole

Fonte: Caselli e Casadio (1995: 138)

Lo scarto tra le due dimensioni dell'attività linguistica comunque permane e prepara ulteriori stadi di incremento lessicale, come si può cogliere nel grafico 4 che mette insieme la comprensione e la produzione nella fase di sviluppo che va da 18-19 mesi a due anni e mezzo. Il ruolo della comprensione appare evidente nel determinare lo sviluppo lessicale. L'osservazione dell'andamento permette di apprezzare una riduzione della distanza tra produzione e comprensione dai 18 ai 30 mesi; tuttavia lo scarto permane, anzi si configura come tratto fisiologico dell'acquisizione lessicale che si mantiene col passare del tempo e statuisce la priorità ineluttabile della ricezione. Dalla ricchezza dell'input protratto nel tempo scaturisce il costituirsi delle condizioni per una produzione linguistica ricca e variata.

Grafico 4 - Comprensione e produzione a confronto



Fonte: Caselli e Casadio (1995: 140)

3.3. *Qualità delle parole*

L'analisi della composizione del vocabolario dei bambini offre elementi per una valutazione qualitativa dei dati quantitativi. Osservando gli studi già citati (in particolare Caselli-Casadio 1995) si trovano conferme a conoscenze già note e indicazioni di nuove evidenze⁴. Considerando l'incidenza media riscontrata in gruppi di parole da 1-50 a >600, si osserva che per ogni gruppo e nel totale complessivo la categoria del nome predomina con una incidenza del 59,05% su verbi, aggettivi, funtori, routines, avverbi. Queste altre categorie sono presenti ma con percentuali di gran lunga inferiori: verbi 14,70%, aggettivi 8,37%, funtori 7,21%, routines 6,80%, avverbi 3,87%. Il rilievo del nome è certo un riflesso dell'attività di denominazione ma è anche un rispecchiamento della predominanza di questa categoria sulle altre nel lessico, come testimoniano ad esempio i dizionari italiani ed inglesi⁵.

Col crescere dell'età, le altre categorie iniziano ad aumentare il loro peso nella composizione del vocabolario a scapito della categoria del nome. In particolare il grafico 5 che raffigura l'incidenza delle specifiche categorie grammaticali nella composizione del vocabolario produttivo permette di cogliere l'incremento graduale delle categorie chiamate a svolgere funzione predicativa, come verbi e aggettivi.

⁴ Le ricerche cui si fa riferimento si fondano sul questionario Mac Arthur, elaborato per la lingua inglese ma riadattato per l'italiano tenendo conto delle sue specifiche peculiarità. Il questionario include una lista di vocaboli su cui i genitori registrano la conoscenza e l'uso che ne fanno i loro bambini in produzione e ricezione. La lista contiene 408 voci ripartite in 19 categorie semantiche per i bambini fino ad un anno e mezzo di età e 670 vocaboli per bambini fino a 30 mesi (cfr. Caselli e Casadio 1995: 45). Le parole sono riconducibili a tre grandi classi: nomi, predicati, funtori.

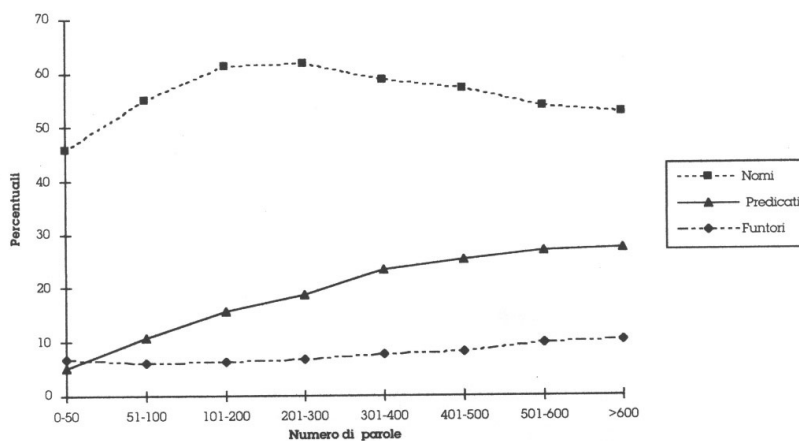
⁵ Si veda ad esempio la categorizzazione dell'intero lemmario del *GRADIT* (De Mauro 2005: 120, tavola 1).

Tabella 3 – Categorie a confronto

| Categorie (%) | Numero di parole in produzione | | | | | | | | Totale |
|--------------------------------|--------------------------------|--------|---------|---------|---------|---------|---------|-------|--------|
| | 1-50 | 51-100 | 101-200 | 201-300 | 301-400 | 401-500 | 501-600 | >600 | |
| Nomi comuni | 29.66 | 42.93 | 52.29 | 52.41 | 49.65 | 47.89 | 45.08 | 44.31 | 51.16 |
| Nomi di persone | 18.04 | 10.22 | 5.95 | 4.37 | 3.62 | 3.59 | 3.74 | 4.27 | 4.76 |
| Suoni della natura | 17.00 | 9.00 | 4.69 | 3.17 | 2.45 | 2.03 | 1.86 | 1.84 | 3.13 |
| Nominali (Nomi+ Persone+Suoni) | 64.70 | 62.15 | 62.93 | 59.95 | 55.72 | 53.51 | 50.68 | 50.42 | 59.05 |
| Routines | 24.10 | 17.82 | 10.25 | 7.61 | 5.78 | 4.62 | 4.10 | 3.94 | 6.80 |
| Verbi | 1.41 | 5.10 | 7.70 | 10.51 | 14.39 | 16.09 | 16.78 | 16.32 | 14.70 |
| Aggettivi | 2.26 | 4.67 | 6.61 | 7.09 | 7.69 | 8.23 | 8.89 | 9.58 | 8.37 |
| Predicati (Verbi+Aggettivi) | 3.67 | 9.77 | 14.31 | 17.60 | 22.08 | 24.32 | 25.67 | 25.90 | 23.07 |
| Funtori | 6.24 | 5.29 | 4.94 | 5.46 | 6.39 | 6.72 | 8.49 | 8.35 | 7.21 |
| Avverbi | 2.26 | 2.73 | 3.10 | 3.22 | 3.40 | 3.68 | 4.32 | 4.18 | 3.87 |

Fonte: Caselli e Casadio (1995: 79)

Grafico 5 – Categorie a confronto



Fonte: Caselli e Casadio (1995: 139)

Il predominio del nome rimane incontrastato e risalta con nettezza anche in lavori in cui si pongono a confronto bambini di lingue diverse. Sulla sua valenza universale però, come si dirà più oltre, vengono espressi dubbi da parte di alcuni studiosi.

La composizione del vocabolario infantile offre elementi di riflessione non privi di interesse anche a fini applicativi. Nel primo gruppo da 1 a 50 parole in bambini con età da 18 a 30 mesi un posto di rilievo è occupato dalle cosiddette formule di *routine*, che rappresentano con un 24,10% una buona fetta della estensione del vocabolario individuale medio. Con *routines* ci si riferisce a quelle espressioni formulaiche, costituite da uno o più vocaboli, relative ad azioni o eventi ritualizzati, come ad esempio *buongiorno/buonanotte, per favore / per piacere, fare il bagno, fare la pipì, pronto, chi è?, cuccù-settete*, la cui acquisizione è determinata dalla forza della funzione pragmatica e dalla reiterazione di questo tipo di scambio comunicativo. La percentuale delle espressioni routinarie prodotte decresce con l'aumentare dell'età e parallelamente all'aumento del numero complessivo di parole.

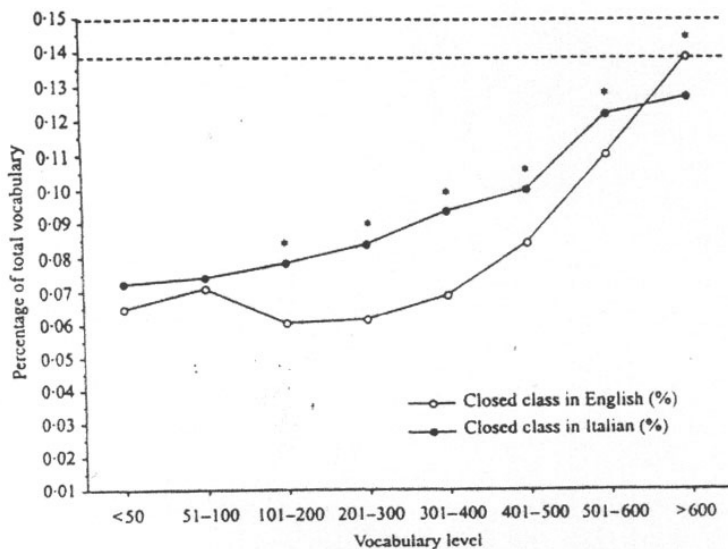
Considerazioni ulteriori riguardanti la qualità della composizione del vocabolario dei bambini italiani emergono nei lavori cross-linguistici italiano-inglese⁶. Tralasciando le molte similarità per le quali si rinvia ai lavori che le illustrano (cfr. Caselli, Bates, Casadio *et al.* 1995; Caselli, Casadio e Bates 1999; Caselli, Casadio e Bates 2001), ci si sofferma su quelle poche evidenze di differenziazione, che

⁶ In Italia e negli USA sono state condotte più inchieste con il medesimo questionario MacArthur. L'identità dello strumento permette analisi e paragoni tra lingue. In particolare, il confronto tra dati relativi all'acquisizione delle prime parole in italiano e in inglese non evidenzia sul piano della quantità elementi di diversità tali da richiedere distinti schemi interpretativi. Vi è un solo elemento dissonante tra le due lingue che emerge dal confronto interlinguistico condotto su gruppi omogenei di bambini diversificati solo per la lingua nativa. I bambini inglesi fanno registrare un maggiore numero di parole rispetto a quelli italiani in quasi tutte le fasce di età, come emerge dallo studio Caselli, Casadio e Bates (1999: 86). L'evidente divario tra i due gruppi linguistici non trova spiegazione. C'è solo da testimoniare che a 1 anno e 6 mesi il punteggio medio del vocabolario si attesta sulle 69 parole per il bambino italiano contro le 114 dell'inglese e a due anni e 6 mesi lo scarto continua a mantenersi con 446 per l'italiano e 534 in inglese. L'apparente 'vantaggio' non è riconducibile a nessuna ragione connessa con la qualità dei campioni o con le condizioni dell'accertamento.

permettono di configurare una qualche ipotesi interpretativa utile per le fasi in cui l'incremento lessicale deve essere sostenuto da attività di insegnamento e apprendimento formali.

Nel confronto fra italiano e inglese, la distribuzione delle categorie segnala l'insorgenza precoce della categoria del verbo nei bimbi italiani rispetto ai loro coetanei anglofoni e un piccolo ma considerevole incremento della classe delle parole funzionali. Non si è in presenza di dati eclatanti: si tratta di spie, di tendenze, che sia sotto il profilo della precocità sia della quantità danno in prima lettura l'impressione di una composizione del vocabolario in cui gli elementi più sensibili alla variazione grammaticale si annuncino comparativamente con maggiore forza rispetto a quella delle classi aperte.

Grafico 6 - Classi chiuse a confronto

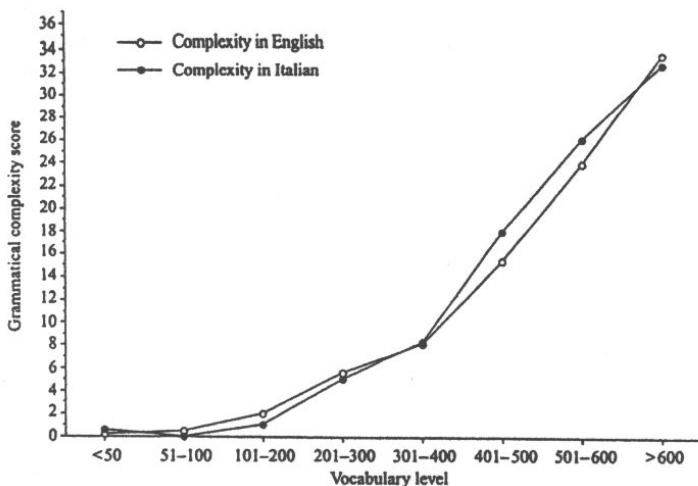


Fonte: Caselli, Casadio e Bates 1999: 94

Ad analisi di dettaglio (cfr. Caselli, Casadio e Bates 1999: 90-108), solo la classe delle parole funzionali si mantiene come tratto di differenziazione rispetto all'inglese, almeno fino alla soglia di >600 parole. Il rilievo che le parole funzionali hanno nel vocabolario viene usato come indice del grado di complessità grammaticale raggiunto dai bambini, indipendentemente dall'età dei soggetti. Considerati dal pun-

to di vista delle ipotesi di sviluppo del vocabolario, il diverso peso delle classi aperte e chiuse suggerisce e conferma una tesi già avanzata e sostenuta in molti lavori dalle stesse autrici, che ci sia “un collegamento potente e forse causale tra sviluppo lessicale e l’emergere della grammatica nel periodo cruciale dello sviluppo” (Caselli, Casadio e Bates 1999: 97)⁷.

Grafico 7 - Complessità sintattica ed estensione del vocabolario



Fonte: Caselli, Casadio e Bates 1999: 97

⁷ L’estensione a lingue diverse ed anche lontane sotto il profilo tipologico è importante quando si cercano indizi funzionali ad una generalizzazione. Talora anche lingue appartenenti alla stessa famiglia possono riservare sorprese. La ricerca sul tedesco (effettuata da Wagner 1985 e citata da Crystal 1993: 244) riporta dati discordanti rispetto a inglese e italiano. Secondo il ricercatore tedesco un bambino di 1 anno e cinque mesi è arrivato a produrre 1.860 tipi di parole diverse e uno di 11 anni oltre 5.000 tipi di parole diverse. Quanto al numero delle occorrenze, rapportando i campioni di parlato a 12 ore si calcola una media di 13.800 occorrenze di parole a un anno e cinque mesi e 25.000 occorrenze a 9 anni e 7 mesi (cfr. Crystal 1993: 244).

3.3. *Universale o locale?*

Il prevalere dei nomi sui verbi ha indotto i ricercatori a considerarlo come un possibile stadio universale nel costituirsi e comporsi del vocabolario individuale (cfr. Caselli, Casadio e Bates 1999: 71), in quanto riflesso di vincoli cognitivi e sociali propri del linguaggio umano. L'ipotesi riscontrata in diverse lingue con studi osservativi e costituitasi inizialmente su campioni di bambini anglofoni è stata messa di recente in discussione da lavori condotti sul cinese e sul coreano. Le analisi cross-linguistiche su bambini cinesi e coreani contraddicono l'universalità del predominio dei nomi ed evidenziano, fin dai primi stadi dell'acquisizione, il prevalere dei verbi su altre categorie. Secondo gli autori di questi studi interlinguistici il prevalere dei nomi sui verbi individuato come carattere universale sarebbe dipendente dalle lingue analizzate e non rispondente a tutte le tipologie di lingue esistenti; si tratterebbe a loro avviso solo di un epifenomeno condizionato essenzialmente dall'inglese, lingua sulla quale è stata messa a punto inizialmente l'ipotesi degli stadi universali di composizione del vocabolario.

Nel discutere i cambiamenti che questi studi sembrerebbero apportare all'ipotesi di un universale dello sviluppo lessicale, Caselli, Casadio, Bates 1999 sostengono che l'italiano potrebbe rappresentare una buona base investigativa per verificare l'insorgere precoce dei verbi. Già la comparazione tra inglese ed italiano ha rilevato l'anticipazione di verbi nel vocabolario dei bambini italiani rispetto a quelli inglesi; inoltre le possibilità di variazione nell'ordine delle parole, le frequenti omissioni del soggetto pronominale determinano nel parlato spontaneo un posizionamento del verbo tale da mettere i bambini italiani in condizioni simili a quelle descritte per il coreano e il cinese e, dunque, nelle migliori condizioni per una verifica del carattere universale nell'acquisizione precoce della categoria del nome.

La verifica dell'ipotesi universalistica si presenta più complessa di quel che appare. Al fondo del divario di interpretazioni non c'è solo una diversa genealogia e tipologia delle lingue ma anche una forte dissimmetria tra i modi e gli oggetti posti a confronto, che dovrebbe a sua volta affrontata e indagata preliminarmente.

4. Il vocabolario medio in età scolare

4.1. Word, word family, parola, lemma

La questione della determinazione dell'oggetto che costituisce l'unità del vocabolario individuale è dirimente per i livelli più avanzati di età, sia per operare raffronti tra studi di una stessa lingua sia ai fini dei confronti interlinguistici. In inglese la *word family* comprende la parola base, le forme flesse della base e le parole derivate dalla base attraverso meccanismi di derivazione trasparenti. L'espressione *word family* rinvia, dunque, ad un insieme di parole legate fra loro da relazioni morfologiche di inequivoca interpretazione. Nel complesso degli studi inglesi sull'ampiezza del vocabolario, l'espressione appare la formulazione più adottata.

Per l'italiano la questione è più complessa. Nella descrizione morfologica dell'italiano non sempre si conviene sulla forma astratta ipotizzata come base di derivazione. Infatti, si adotta tanto una forma astratta costituita dalla radice lessicale e dalla vocale tematica, come ad esempio *canta*, da non confondere con la terza persona singolare dell'indicativo presente, quanto una delle sue forme attualizzate, come ad es. l'infinito *cantare*, adattando di conseguenza le regole per la generazione di altre parole. Operando con famiglie di parole, la ricca morfologia flessiva fa esplodere le serie di parole correlate. Anzi, comprendendo parole base, flesse e derivati trasparenti sotto un'unica entrata, il numero effettivo delle parole conosciute è quasi impossibile da determinare. In italiano ad esempio, *giocare* può contare da solo come un elemento o, se lo si considera come esponente dell'intera famiglia lessicale, può rappresentare un insieme difficilmente quantificabile o determinabile solo in via approssimativa. Esso infatti deve inglobare forme flesse per ogni tempo verbale semplice, forme per ciascun tempo infinito, tutti i derivati con i suffissi, ad esempio *-tore*, *-trice*, *-ata*, *-0* (suffisso zero), *-ino*, *-etto*, *-erello*, *-abile*, *-erellare*, *-icchiare*, *-oso*, ecc., le doppie derivazioni *gioc+at(a)+ona*, *giocatina*, *giocherello*, *giocherellone* con tutte le variazioni di forma a cui la morfologia le assoggetta. *Giocare* può rappresentare solo se stesso

oppure, come conseguenza della flessione, delle possibilità di derivazione e delle combinazioni morfologiche, essere il capostipite di una famiglia di almeno 72 ‘parole’. Il numero delle parole derivate è certamente inferiore a quello registrato in qualunque dizionario dell’italiano e a quello delle parole potenziali. Inoltre ad esso andrebbero sommate le espressioni polirematiche formate con *giocare*, quali *giocare a carte scoperte*, *giocare d’astuzia*, *giocare con le/sulle parole*, *giocare tutte le carte*, *giocare in borsa*, *giocare sul velluto*, *giocare di sponda*, *giocare al rialzo* ecc.

Due modi diversi di intendere ‘parola’, come unità singola o come insieme di parole danno ovviamente risultati diversi. I lavori di quantificazione dell’italiano, diversamente dall’inglese, optano per fare riferimento alle entrate del dizionario, ai lemmi. Seguendo questa indicazione, i derivati attestati con significati definiti hanno entrata propria: ciascun derivato conta, dunque, per una parola e le forme flesse non vengono computate in quanto derivabili dall’applicazione delle regole morfologiche della lingua e non direttamente imputabili al lessico. Questo orientamento permette di superare le differenze tra le morfologie di lingue diverse, che rendono ardui i confronti interlinguistici soprattutto in sistemi in cui le scelte flessionali operano entro un diverso e rilevante numero di opzioni. Nella stessa direzione sembra orientarsi Aitchison (2003) nel suo lavoro sul lessico mentale.

An educated adult speaker of English can understand, and potentially use, at least 50000 words, with a word provisionally defined as ‘a dictionary entry’. Modern dictionaries usually include different forms of a word under the same entry, so *sing*, *sings*, *sang*, *sung* would all come under the headword *sing*. However, they normally provide separate entries for derivatives whose meaning cannot be reliably guessed, so *singer* would have an entry to itself, because it does not just mean ‘someone who sings’, but more usually ‘someone who sings for a living’ (Aitchison, 2003: 7).

Molti studi italiani sull’estensione del vocabolario medio dei parlanti si orientano su questa stessa strada e misurano l’ampiezza della conoscenza rapportandola alle entrate del dizionario sia quello di base per misurare la conoscenza del lessico fondamentale sia ad un dizionario di medie dimensioni per saggiare il grado complessivo di conoscenza del lessico.

4.2. *Quantità e qualità in inglese*

Molti lavori di area anglosassone sull'insegnamento e apprendimento del vocabolario (cfr. Nation 1990 e 2001; Nation e Waring 1997; Schmitt 2000; Schmitt e McCarthy 1997; Stahl 1999) riportano dati sull'entità delle unità lessicali conosciute da parlanti inglese come lingua nativa. Studi più generali sul lessico mentale (cfr. Aitchison 2003³; Singleton 1999) fanno riferimento a ricerche che forniscono quantificazioni del numero di parole conosciute da soggetti ventenni con grado medio di istruzione. Nelle due aree consultate i dati non sono di prima mano; essi rinviano a ricerche talora datate e mettono insieme informazioni di fonti diverse senza fare cenno ai criteri metodologici adottati nel corso delle rilevazioni. Alcune fonti ripropongono gli stessi numeri e le stesse citazioni; altre forniscono dati in dissonanza senza giustificare la congruità/incongruità delle cifre.

Nation sottolinea che il tratto saliente delle stime riguardanti l'estensione del vocabolario di un parlante inglese nativo sia la *enormous variation in the estimates* (1990: 11). La variazione si giustifica da un canto con l'ambiguità del termine 'parola' di cui si è detto anche riguardo alla numerosità delle entrate dei dizionario, dall'altro con le difficoltà intrinseche di una osservazione diretta che, dopo i primi anni in cui si opera con piccoli numeri a tre cifre, si offre al ricercatore con estensioni difficilmente controllabili. Nation condivide le stime di quanti attribuiscono ad uno studente universitario la conoscenza di circa 20000 parole (*words*), con una ipotesi progressiva di acquisizione e/o apprendimento di circa 1.000 parole annue, o da 3 a 7 parole al giorno. Lo studioso, collazionando fonti diverse, riassume in una tabella i dati sullo sviluppo lessicale, in cui la progressione incrementale viene valutata dal primo anno di vita fino ai 18 anni di età.

Tabella 4 - Estensione del vocabolario medio di un parlante inglese

| <i>Fonte</i> | <i>N. parole</i> | <i>Età</i> |
|-------------------|------------------|----------------|
| Kirkpatrick | 235 | 1 anno 3 mesi |
| Kirkpatrick | 400 | 2 anni 8 mesi |
| Kirkpatrick | 700 | 3 anni 8 mesi |
| Salisbury | 1.528 | 5 anni 5 mesi |
| Termon and Childs | 2.500 | 6 anni 5 mesi |
| Kirkpatrick | 4.480 | 8 anni 5 mesi |
| Kirkpatrick | 6.620 | 9 anni 6 mesi |
| Kirkpatrick | 7.020 | 10 anni 7 mesi |
| Kirkpatrick | 7.860 | 11 anni 7 mesi |
| Kirkpatrick | 8.700 | 12 anni 8 mesi |
| Kirkpatrick | 10.660 | 13 anni 9 mesi |
| Kirkpatrick | 12.000 | 15 anni |
| Kirkpatrick | 17.600 | 18 anni |

Fonte: Nation 1990: 12

Nella tabella costruita con il contributo di più ricerche, l'ampliamento del vocabolario è sì costante ma ha accelerazioni o incrementi più contenuti a seconda di delicate fasi dello sviluppo. Per quanto straordinario, l'incremento lessicale è lontano dal rapportarsi alle rappresentazioni del lessico che offrono i dizionari inglesi. Un adulto nativo, a conclusione dell'intero ciclo di istruzione, avrebbe un patrimonio lessicale che rappresenterebbe meno di un quinto di un dizionario tarato per le necessità degli studenti, come ad esempio il Cambridge.

La determinazione delle stime è lontana dall'essere univoca. Anche operando con parole, Aitchinson (2003) presume una espansione di 54.000 parole diverse per un adulto istruito; Barrett (1999: 244) sostiene che il campo di variazione entro cui oscillano le conoscenze di

un ventenne istruito va da 50.000 a 250.000 *words*⁸; in questo ampio spettro di possibilità, egli si attesta su *the mid-figure of 150000 words*. Dall'una all'altra ricerca corre un tale divario da rendere vano lo sforzo di circoscrivere in quantità il vocabolario medio di un parlante inglese.

Altri studiosi anglosassoni ritengono che le cifre relative all'estensione del vocabolario dovrebbero essere intese estensivamente: esse, infatti, non sarebbero da riferirsi a singole parole bensì a famiglie di parole. Con questa interpretazione estensiva, la Tabella 7 diviene compatibile con la stima di circa 20000 *word families* che dovrebbe corrispondere, secondo molti studiosi, all'estensione di un *university graduate* nativo (cfr. Schmitt 2000: 3). Da questa cifra globale riferita ad un adulto di circa 20 anni, Nation e Waring traggono la conseguenza che "*native speakers will add roughly 1000 word families a year to their vocabulary size. This means that a five year old beginning school will have a vocabulary of around 4000 to 5000 word families*" (Nation e Waring, 1997: 7-8). Questa divisione di indubbia comodità rappresentativa nasconde di fatto le accelerazioni e gli stalli che sembrano contraddistinguere il costituirsi del vocabolario individuale nei primi 20 anni, a favore di un aumento continuo, regolare e costante che è tutto da dimostrare. Essa inoltre proietta su un bambino di sei anni un patrimonio lessicale di gran lunga superiore alle effettive 2.500 parole attestate da altre ricerche (cfr. Tabella 5). La distanza tra le 2.500 attestate e le 6.000 parole presunte dà il senso dell'approssimazione delle stime e della loro bassa attendibilità. Dal

⁸ Secondo quanto Barrett stesso ha esplicitato su richiesta, le sue indicazioni si fondano su misurazioni condotte e pubblicate da Seashore and Eckerson nel 1940. Nel loro ambito, come da appunto fornito da Barrett, *word* "means the headwords listed in Funk and Wagnall's Dictionary" e, dunque, suscettibili di ulteriori espansioni quantitative.

punto di vista dell'apprendimento (e dell'insegnamento), lo scenario cambia radicalmente considerando l'una o l'altra stima lessicale e intendendo riferire i numeri a *word* o a *word families*.

4.3. *Quantità e qualità in italiano*

Molti studi italiani sull'estensione del vocabolario medio propendono per la misurazione di parole e non di famiglie di parole. Così facendo, i ricercatori riescono a rapportarsi alle entrate di un dizionario italiano di medie dimensioni e a valutare l'estensione delle conoscenze individuali misurandola rispetto al lessico rappresentato dallo strumento lessicografico. Si raggiungono con questa strategia due scopi distinti, la quantificazione del vocabolario individuale medio e lo scarto di questo rispetto alla rappresentazione dizionaristica del lessico italiano e alla sua stratificazione.

L'estensione del vocabolario italiano nella fascia dai 6 ai 18 anni non è in Italia oggetto di indagine campionarie sistematiche né nel tempo né nello spazio, che spesso, come è accaduto con le indagini sulla lettura, rivela ampi divari nelle conoscenze e nelle abilità tra Nord, Centro e Sud ed anche tra Est ed Ovest⁹. Come per i piccoli, così per la popolazione scolastica i dati a disposizione sono parziali e non riconducibili all'intera popolazione. Si tratta di lavori su singole realtà che però forniscono elementi utili per dare un'idea, anche se in via approssimativa, della effettiva estensione delle conoscenze lessicali a diversi livelli di età ed istruzione. Si osservino i dati a disposizione.

In alcuni lavori di area psicologica si trovano indicazioni numeriche che non permettono di risalire alle fonti per controllarne metodo-

⁹ Cfr. Oltre alle inchieste internazionali condotte e pubblicate dalla *International Association for the Evaluation of Educational Achievement* (IEA), consultabili sul sito <http://www.iea.nl/iea/hq/>, per un'analisi dei dati italiani si veda Lucisano 1994.

logia e congruità. Due dati in particolare attraggono per le cifre. L'uno riportato da Oliverio (2001) che attribuisce ad un quindicenne una oscillazione nel vocabolario tra le 10.000 e 30.000 parole: la forbice appare troppo ampia per essere assunta come valore medio della conoscenza lessicale adolescenziale. L'altro dato (cfr. Caselli e Volterra 1999) stima per un diciottenne alla fine delle scuole superiori una consistenza lessicale pari a circa 40.000 parole. Anche in questo caso non è facile risalire alla fonte della quantificazione e, dunque, l'attribuzione seppure congruente con altri dati parziali non può essere assunta come base attendibile. Per avere cifre controllabili bisogna affidarsi a lavori parziali che offrano garanzie nella metodica dell'accertamento.

Un recente lavoro offre uno spaccato su comprensione e produzione di vocaboli da parte di studenti tredicenni con otto anni di istruzione (cfr. Tomassi 2003). La metodologia seguita – rilevazione su lista di vocaboli – la scelta e il controllo del test, la taratura permettono di assumere questo lavoro come base di discussione per valutare quantitativamente l'entità del vocabolario degli adolescenti. Il lavoro, per le proiezioni stratificate sulla composizione del lessico, ammette anche, pur nella parzialità del campione, una valutazione qualitativa.

La media dell'estensione del lessico si aggira intorno alle 38.000 parole, considerando unitariamente *le parole capite non usate* e *capite e usate*. Per completezza, si riportano anche la quantificazione di tutte le altre possibilità ammesse dalle risposte al questionario lessicale: le parole *mai sentite* vanno dalle 30.000 alle 35.000; le parole *sentite non conosciute* vanno dalle 25.000 alle 30.000; le parole *capite non usate* non superano le 5.000. Il dizionario adoperato¹⁰ come base campiona-

¹⁰ Il dizionario della Paravia in volume unico (cfr. De Mauro 2000) contiene 140.000 lemmi e 30.000 espressioni polirematiche. Ogni entrata del dizionario contrassegna i suoi lemmi, semplici o polirematici, con una marca che ne specifica il livello d'uso e l'ambito di appartenenza. Le marche adoperate specificano per ogni en-

ria adotta la marcatura delle entrate lessicali. Ciò vuol dire che ogni lemma è riconducibile ad una specifica fascia della stratificata composizione del lessico italiano. Il vocabolario di base (VdB), nucleo centrale del lessico che innerva ogni enunciato, risulta conosciuto per il 90,18% dal primo gruppo e per il 79,76% dal secondo gruppo, con una forbice in cui coesistono un individuo che centra il 100% del riconoscimento e un altro soggetto con appena il 21,43%. La variazione individuale emersa già nelle prime fasce di età sembra permanere anche tra gli adolescenti, ma va detto che solo pochi casi si collocano agli estremi del minimo e massimo delle risposte mentre la maggioranza esibisce in massa la conoscenza del nucleo più solido del VdB. Le parole fondamentali note al 94,87% sono condivise da una fascia di studenti che va dal 75 al 100%. Il cuore primigenio della lingua italiana, già attestato in Dante e presente in ogni discorso o testo in percentuali che vanno oltre il 94%, è patrimonio della stragrande maggioranza dei soggetti intervistati.

Esistono anche per la fascia di lessemi del vocabolario di base lacune imprevedibili e impreviste, soprattutto se si tiene conto delle previsioni di espansione lessicale che fissano come obiettivo in uscita dalla scuola di base dopo otto anni di scolarità proprio l'intero VdB (cfr. Gensini, Vedovelli 1983). A parziale giustificazione della lacune lessicali emerse nel pur piccolo nucleo di vocaboli fondamentali, va detto che nessun programma scolastico ufficiale ha posto il lessico tra gli obiettivi da raggiungere limitando gli scarni accenni alla necessità di un generico ampliamento.

trata se si tratta di vocabolo fondamentale (FO), di alto uso (AU), di alta disponibilità (AD), comune (CO), tecnico-specialistico (TS), di uso esclusivamente letterario (LE), dialettale (DI), regionale (RE), esotismo (ES), di basso uso (BU), obsoleto (OB). Le parole fondamentali, di alto uso e di alta disponibilità costituiscono il vocabolario di base della lingua italiana; si tratta di un insieme di poco meno di 7000 parole che copre con le sue occorrenze una percentuale tra il 94 e il 96% di qualsiasi discorso o testo prodotto in italiano.

Se si spinge l'analisi qualitativa fino al settore del dizionario marcato 'comune' – si tratta di circa 40.000 vocaboli su 140.000 – esso risulta conosciuto nel primo gruppo per il 49,50% e al 35,17% nel secondo. Non è corretto per l'esiguità del campione riportare tali dati all'intera popolazione studentesca; esso però manifesta una tendenza ad andare ben oltre i confini dell'indispensabile vocabolario di base e a cogliere nelle possibilità di variazione sinonimica con i vocaboli comuni e con gli altri strati lessicali un elemento di crescita personale e culturale non banale. Un ultimo dato tratto da un diverso lavoro (Tomassi 2003) può contribuire a dare una idea delle capacità di incremento autonomo delle conoscenze lessicali dei giovani adolescenti rafforzando l'ipotesi di una correlazione funzionale tra la complessità dei processi grammaticali e l'ampiezza del vocabolario. Prefissi, suffissi e confissi, considerati entrate autonome dal dizionario adottato come sfondo e in quanto tali inserite nella campionatura dei vocaboli, sono stati riconosciuti da circa il 30% dei giovani intervistati. Ciò vuol dire che per questa percentuale di studenti, attraverso la gestione di elementi salienti nei processi di derivazione e composizione, può generarsi in modo autonomo la comprensione e la produzione delle serie derivative e dei composti lessicali.

Il livello universitario soffre della stessa frammentarietà e dispersione dei dati quantitativi. Le ricerche censite sono relative a rilevazioni su gruppi casuali di universitari non rappresentativi dell'intero universo. Prive della rappresentatività necessaria a consentire di proiettare i risultati ottenuti sull'intera popolazione studentesca, le inchieste vanno lette come indicazione tendenziali che meriterebbero inchieste campionarie controllate e reiterate nel tempo. Le due indagini cui si fa riferimento in questo lavoro (cfr. Boni 1982; Franciotti 2003) hanno punti comuni e differenze: sono distanti temporalmente, essendo state condotte l'una nel 1981 e la seconda nel 2002; operano su check list costruite sulla base del stesso dizionario della lingua italiana in due diverse edizioni (Zingarelli 1980 e 2000); adottano l'autovalutazione dei soggetti intervistati. Le similarità di metodo e somministrazione non sono sufficienti a far trarre conseguenze riguardo alle differenze che pure si riscontrano nei risultati. Più che il confronto è utile sintetizzare i dati raccolti nella prima inchiesta del 1982.

L'indagine Boni 1982 attesta che le 1.000 parole della lista risultano conosciute al 68,2%. Il dato rapportato alle cifre assolute sta ad indicare che gli universitari in uscita nell'anno accademico 1981-1982 conoscevano un po' più di 70.000 parole delle circa 100.000 presenti nello Zingarelli 1980. Il risultato ottenuto nella indagine supera di circa 10 punti la previsione assunta nell'ipotesi di partenza che prevedeva per un parlante colto una quota di vocabolario pari al 60% di un lemmario medio¹¹.

Nella gamma delle risposte universitarie si registra il permanere di quella variabilità individuale che è stata rilevato anche nelle inchieste sugli adolescenti; l'entità della variazione non inficia però l'attendibilità del dato medio. Connessa ad uno stile di apprendimento, ad una inclinazione personale, al variare dello spessore culturale l'oscillazione nell'estensione del vocabolario risente del tipo di facoltà e soprattutto del tipo di diploma medio-superiore conseguito. La distribuzione delle risposte in base alle facoltà indica che le facoltà umanistiche sopravanzano quelle scientifiche in modo non clamoroso ma con una distanza apprezzabile di qualche punto percentuale; le facoltà scientifiche di contro prevalgono nel dominare meglio il lessico specialistico¹². Si tratta di dati interessanti per osservare caratteristiche e tratti del vocabolario medio del parlante colto e per trarne indicazioni per interventi lessicali mirati alla riqualificazione della conoscenza lessicale.

¹¹ A venti anni di distanza, Franciotti 2003 rileva una percentuale media leggermente inferiore, pari al 65,7% e più vicina all'ipotesi postulata nel 1982 per il parlante colto.

¹² Come chiusa si riportano due elementi di curiosità su cui riflettere: gli studenti della facoltà di Giurisprudenza hanno una estensione lessicale che sopravanza quella di tutti gli altri intervistati (cfr. Boni 1982, ma anche Franciotti 2003); gli universitari di Scienze della comunicazione, in controtendenza rispetto alla specificità dei loro studi, risultano gli studenti con un vocabolario medio più basso in assoluto (cfr. Franciotti 2003).

5. Conclusioni per un rilancio

L'intervallo di tempo che le ricerche sul vocabolario individuale degli italofoeni coprono è abbastanza vasto. Anche le metodologie adottate e le parzialità dei soggetti indagati oscillano in un ampio spettro di alternative possibili, tanto che di alcune ricerche non si è fatta menzione per la difficoltà di porle in comparazione con quelle esibite. La diversità complessiva che caratterizza questo settore della ricerca lessicale richiede che si proceda con regolarità e uniformità metodologica alla rilevazione sistematica delle capacità lessicali medie in ogni ordine e grado di istruzione e nella popolazione adulta. Ciò che si ricava dallo spoglio di dati è, dunque, la necessità di disporre di indagini campionarie longitudinali in grado di dar conto di eventuali variazioni, contrazioni o espansioni, del patrimonio lessicale medio con qualche attendibilità e certezza. Il rilevamento assumerebbe senso se inserito in un piano di alfabetizzazione lessicale da avviare nelle scuole, che risponda all'istanza di migliorare l'input per rafforzare e favorire l'acquisizione spontanea e si prefigga al contempo di estendere in quantità e ampliare in profondità le conoscenze lessicali con programmi mirati di insegnamento/apprendimento.

Si tocca a questo riguardo una questione controversa e delicata relativa alla competenza lessicale, che non ha trovato finora adeguata sistemazione nell'ambito della linguistica educativa. Diversamente dalla tradizione teorica di impianto generativo in cui il termine competenza è stato elaborato, gli studi applicativi fanno un uso estensivo del termine.

Chomsky a utilisé la notion de compétence dans un sens que connaissent les linguistes et les didacticiens de la langue et auquel certains d'entre eux restent fidèles. Dès lors, le sens que prend ce concept dans le champ du travail et de la formation leur paraît une déviation, une source de confusion, voire une aberration théorique. Le premier problème naît du fait qu'on utilise le même mot pour désigner des *réalités différentes*, quand bien même elles présentent une part de signification commune. Aucune acception n'est *a priori* plus légitime ou plus absurde que l'autre, d'un point de vue sémantique. La question est de savoir ni [si] nous voulons nommer compétence une faculté innée de l'espèce humaine ou un *acquis* qui appartient à chaque personne et varie donc de l'une à l'autre (Perrenoud 2001 : 17).

Seguendo l'indicazione francese ma anche l'uso estensivo che ne propone Lavinio¹³, con il termine competenza si può intendere tanto la conoscenza della lingua che è innata per aspetti molto generali quanto l'elaborazione frutto di acquisizione e di apprendimento anche formale, variabile da un individuo ad un altro in ragione delle esperienze anche formative.

Nel caso delle conoscenze lessicali l'apporto dell'insegnamento e dell'apprendimento gioca un ruolo non secondario, eppure di rado esso è preso in considerazione nella definizione della competenza lessicale. Ampliando l'orizzonte, più in generale si può affermare che è il lessico nel suo complesso a rimanere in ombra nelle teorie esplicative della conoscenza linguistica. Come è noto, infatti, entro i confini di teorie elaborate per la sintassi il ruolo del lessico è marginale, anche se nel tempo qualcosa si è modificato a favore di una sua maggiore considerazione (cfr. Chomsky 1995: 30-33 e 235-241). Studiosi diversamente orientati hanno talora enfatizzato il ruolo del lessico per distanziarsi dalle teorie generative ortodosse (cfr. Jackendoff 2002: 152-195) senza con ciò pervenire ad una elaborazione teorica complessiva, entro cui il lessico trovi una sua adeguata esplicazione. Le questioni connesse al ruolo e al significato da attribuire al lessico nella teoria anche solo glottodidattica sono troppo grandi per essere affrontate in questa sede; tuttavia è condivisibile ed applicabile all'intero lessico quanto afferma Jackendoff per il senso.

I am interested in constructing a stance on meaning from which it is possible to make sense of the sort of detailed empirical investigation that linguists do. The relation between the philosophy and the dirty work has to be a two-way street. (Jackendoff, 2002:268)

¹³ In questo senso la definizione formulata da Lavinio in Lavinio e Vedovelli 1997, s.v. Diversamente da Perrenoud 2001 ma in consonanza con Spolsky 1999, in altra sede si è tentato di affermare l'autonomia della linguistica educativa (De Mauro-Ferreri, in stampa), nel cui ambito può trovare formulazione una concezione di competenza che inglobi gli esiti di acquisizione e apprendimento.

Le elaborazioni su e intorno al lessico devono operare in due direzioni. La formulazione di ipotesi teoriche può rappresentare una delle direttrici di marcia della strada a doppio senso di cui parla Jackendoff; sull'altra deve viaggiare il dato 'sporco' rappresentato dall'estensione del vocabolario individuale con il diverso carico di conoscenza qualitativa e quantitativa: solo i due sensi di marcia insieme fanno una *two-way street*. Pur non avendo considerato in questo lavoro gli aspetti qualitativi del conoscere le parole – altra medaglia della competenza lessicale (cfr. Nation 1990: 31; Laufer 1997: 141; Schmitt 2000: 5) – il solo valore numerico del vocabolario individuale (frutto di ampliamento progressivo e continuo ma nel tempo soggetto a possibili contrazioni e riduzioni) rappresenta, cifre alla mano, una buona base dimostrativa della elaborazione mai conclusa della componente lessicale. La posizione da assumere riguardo al lessico deve tenere in conto l'estrema variazione delle conoscenze individuali in quantità e qualità.

La costruzione delle conoscenze lessicali oltrepassa i vincoli dell'acquisizione spontanea e si costituisce nel corso del tempo sulla base di input mirati volti ad attivare tutta la gamma degli usi linguistici. Sembra nascondersi nell'allargamento dello spazio linguistico e culturale, nell'esercizio intelligente dell'uso linguistico riflessivo la chiave di accesso all'ampliamento del vocabolario e al costituirsi della profondità della sua conoscenza. I due aspetti della quantità e della qualità non possono mai essere scissi. Un piano di misurazione delle effettive capacità ricettive e produttive del lessico della propria lingua materna non può prescindere dal prevederli insieme; e un progetto di espansione delle capacità lessicali deve perseguire ampiezza e profondità in modo da innescare un circuito virtuoso capace di generare per autoapprendimento un incremento continuo.

Un piano di insegnamento / apprendimento del lessico deve concorrere ad agevolare la conservazione delle conoscenze lessicali nella memoria a lungo termine favorendo l'acquisizione di quegli elementi di qualità che facilitano i meccanismi di rappresentazione, immagazzinaggio e accesso al lessico. Tra i molti elementi che entrano in un piano di alfabetizzazione lessicale (cfr. Ferreri, *Parole tra quantità e qualità*, in stampa), particolare attenzione è riservata agli aspetti in grado di rinforzare i meccanismi acquisizionali già esistenti: si favoriscono i rapporti di senso per specificare per differenza e confronto il

significato e le accezioni delle parole; si lavora sulla configurazione sintattica e le collocazioni per dare il senso delle solidarietà sintagmatiche; si attiva il riconoscimento e la produzione degli affissi, sia prefissi sia suffissi, e dei confissi (elementi come *biblio-*, *crono-*, *-fono-* ecc.) per generare le regole di funzionamento di derivazione e composizione: tutti elementi che fanno assieme quantità e qualità della conoscenza lessicale e potenziale di sviluppo per un apprendimento che sappia autogenerarsi di continuo per le necessità della *parole* e della lingua.

Su un aspetto cardinale vale la pena soffermarsi. Al primo posto di un piano di alfabetizzazione lessicale è opportuno porre il vocabolario di base e, *in primis*, il suo nucleo di 2.000 parole fondamentali. Le ragioni a base di un'insistenza mai sufficiente sull'importanza del VdB e delle sue prime 2.000 parole sono di diverso ordine. Come è ben documentato per l'italiano e per l'inglese, a fronte di molte centinaia di migliaia di occorrenze prodotte dai parlanti nelle più varie occasioni le parole diverse sono di gran lunga inferiori. La tabella elaborata sull'inglese (Tabella 8) su una base più larga dei conti fatti per l'italiano nei lessici di frequenza permette di osservare con estrema chiarezza il contrarsi dei numeri nel passaggio da occorrenze a parole; è, inoltre, significativa la percentuale di penetrazione che le parole più frequenti hanno sull'intero corpus di cinque milioni di occorrenze.

Tabella 5 - Conteggio effettuato su 5 milioni di occorrenze di parole inglesi

| <i>Parole</i> | <i>% di occorrenze</i> |
|---------------|------------------------|
| 10 | 23,7 |
| 100 | 49 |
| 2.000 | 81,3 |
| 3.000 | 85,2 |
| 5.000 | 89,4 |
| 43.831 | 99 |
| 86.741 | 100 |

Fonte: Nation 1990: 17, rielaborata)

Sul versante didattico queste cifre sono di indubbia significatività. Cinque milioni di occorrenze sono repliche di circa 87.000 parole diverse; di queste migliaia di parole, le 10 più frequenti si replicano fino

a costituire quasi il 24% di 5 milioni: un milione e duecentomila occorrenze di parole sono repliche di dieci parole; le 100 parole più usate rappresentano il 49 per cento del totale, quasi due milioni e cinquecentomila occorrenze.

Alle prime 2.000 parole più frequenti è riservato il compito di innervare con percentuali altissime ogni discorso o testo prodotto in italiano o in inglese. Per l'italiano si tenga presente che, con oscillazioni di circa 8 punti percentuali legati all'oggetto di discorso e alle condizioni di produzione (situazione, destinatario, mezzo di comunicazione, tipo di testo, scopo), le parole fondamentali rappresentano nelle loro varie forme il 94% di qualsiasi discorso o testo si ascolti o si legga. Appare indubbia l'utilità di garantire ad ogni allievo il possesso attivo di questo nucleo lessicale.

Agli aspetti di interesse prevalentemente didattico del VdB si affiancano questioni di portata più generale che riguardano l'agevolazione del costituirsi della conoscenza lessicale. L'ulteriore significato assunto dal VdB nel piano di alfabetizzazione lessicale si coglie meglio se si leggono le funzioni e le applicazioni che Martin (1989) indica per i dizionari di frequenza.

Next to their purely linguistic descriptive function frequency dictionaries are often used by *language didacticians* in their attempts to construct a basic vocabulary. Frequency dictionaries are also research instruments for *psycholinguistics* where they provide material for all kind of experiments (e.g. word recognition tests) or can bring evidence for the construction of the mental lexicon. Furthermore one will find frequency dictionaries used in *stylometric studies* mostly as a kind of norm or yardstick against which possible deviations can be evaluated. Last but not least a well conceived frequency dictionary can lend excellent services in the construction of the lexical component of *natural language processing systems* Martin (1989: 1319).

Pare emergere da queste considerazioni ancora di più di quanto l'autore non dica. Gli psicolinguisti che mettono a punto test per la ricognizione di parole, gli informatici che simulano su macchine il funzionamento del linguaggio naturale necessitano per i loro scopi di parole ordinate in base alla frequenza. Le necessità dei ricercatori esaltano un'istanza che è propria dei parlanti¹⁴. Per tutti i locutori le parole più usate sono di primaria utilità al di là di ciò che esse significano e oltre ciò che per se stesse trasmettono. Per consentire l'elaborazione dei significati, per permettere la rappresentazione del lessico mentale, le parole richiedono altre parole. Per costruire nuovi significati i parlanti ricorrono alle parole più radicate nella coscienza individuale e collettiva, a quelle parole che, come esito dell'attività locutoria dell'intera comunità sedimentatasi attraverso il tempo, appaiono di immediata disponibilità. Assicurare il pieno possesso del vocabolario di base di una lingua è compito non eludibile di qualsiasi progetto di apprendimento; esso diviene essenziale nell'alfabetizzazione lessicale per consentire a tutti il pieno dispiegarsi del potenziale dei significati.

¹⁴ Sul ruolo della frequenza d'uso nel riconoscimento delle parole cfr. Colombo 1993, ma anche Aitchinson 2003. Per la significatività della frequenza nei processi di rappresentazione della struttura morfologica cfr. Laudanna e Burani 1999.

