



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Università degli Studi di Padova

Dipartimento di Studi Linguistici e Letterari

Corso di Laurea Magistrale in Linguistica
Classe LM-39

Tesi di Laurea

Le competenze lessicali negli ultimi anni della scuola primaria. Un'indagine sul lessico della cittadinanza

Relatore
Prof.ssa Maria Teresa Vigolo

Correlatore
Prof.ssa Loredana Corrà

Laureanda
Giulia Colesso
n° matr.1082163 / LMLIN

Anno Accademico 2014 / 2015

I limiti del mio linguaggio significano i limiti del mio mondo.

Ludwig Wittgenstein

Indice

Introduzione	1
1. Il lessico	3
1.1. Che cos'è il lessico	3
1.1.1. Lessico, vocabolario, dizionario	3
1.1.2. Lessicologia e lessicografia	5
1.1.3. La lessicalizzazione	6
1.2. Che cosa sono le parole	8
1.2.1. Tipi di parole	9
1.3. La stratificazione del lessico	12
1.3.1. Stratificazione per origine	12
1.3.2. Stratificazione per frequenza d'uso	13
1.4. Significato, semantica e campo semantico	15
1.5. Il lessico mentale	16
1.6. I dizionari di frequenza della lingua italiana e i dizionari di nuova generazione	18
1.7. Il vocabolario di base	20
1.7.1. Come è stato realizzato	21
1.7.2. Dati quantitativi del VdB	23
1.8. Il Lessico Elementare	26
2. Lo sviluppo lessicale nei bambini e come misurarlo	32
2.1. Cenni sulle teorie di acquisizione del linguaggio	32
2.2. Lo sviluppo lessicale nella terza infanzia	34
2.3. Gli obiettivi stabiliti dal MIUR	36
2.4. I test lessicali	40
2.5. Le capacità definitorie dei bambini	43
2.6. Ricerche sulle competenze lessicali dei bambini	45
3. La ricerca sperimentale	51
3.1. Le scelte preliminari	51
3.2. I questionari	52
3.3. La raccolta dei dati	53
3.4. Osservazioni durante la somministrazione	54

3.5. Limiti della ricerca e possibili sviluppi	55
3.6. Analisi quantitativa dei dati	56
3.7. Le definizioni	63
3.8. Riflessioni conclusive	67
Conclusioni	69
Bibliografia	71
Appendice 1	75
Appendice 2	76
Appendice 3	81

Introduzione

Quali parole conoscono i bambini? Il significato che attribuiscono alle parole quanto si avvicina a quello normalmente condiviso con il resto della comunità linguistica? Da questi interrogativi è nata la curiosità di indagare il lessico in una particolare categoria di parlanti: i bambini degli ultimi anni delle elementari.

In letteratura i contributi dei linguisti si concentrano soprattutto sul linguaggio degli adulti, quelli degli psicologi sui processi d'acquisizione linguistica nei neonati, mentre la fascia d'età comprendente bambini e adolescenti è stata per lungo tempo trascurata da entrambe le discipline, lasciando un terreno fertile a nuove ricerche. E ancor oggi, nonostante dagli anni Ottanta siano aumentati i contributi in quest'ambito, ci sono molti versanti inesplorati: i contributi di maggior rilievo provengono soprattutto dalla psicolinguistica e da ricercatori che si occupano di educazione linguistica e glottodidattica.

Con questo lavoro si vuole portare un piccolo contributo nello studio del lessico in questa fascia d'età. Si è cercato, tramite una ricerca sul campo, di capire quali parole del vocabolario di base (De Mauro 2003) dell'area semantica della cittadinanza sono conosciute dai bambini di terza, quarta e quinta primaria e quale idea hanno questi bambini delle parole che dichiarano di conoscere. Si è partiti, infatti, dal presupposto che le conoscenze lessicali sono un tassello fondamentale della competenza linguistica: il lessico è quell'aspetto di una lingua che più degli altri permette di guardare alla società e alla cultura di cui la lingua è espressione.

Nella presente tesi si tratterà innanzitutto del lessico, cercando di capire come è strutturato, e delle parole. Si analizzeranno poi le stratificazioni interne del lessico e si accennerà ai principali strumenti lessicografici di nuova generazione, con particolare riguardo al Lessico Elementare, individuato nel 1994 da Lucia Marconi e colleghi, e al vocabolario di base, di cui si è occupato in primo luogo Tullio De Mauro.

Nel capitolo successivo si accennerà alle diverse teorie esistenti sull'acquisizione del linguaggio e si andrà poi a focalizzare l'attenzione sullo sviluppo del lessico

durante la cosiddetta terza infanzia. Di quest'età si parlerà anche prendendo in considerazione le disposizioni date dal Ministero dell'Istruzione, dell'Università e della Ricerca, in riferimento in particolare agli obiettivi scolastici da raggiungere durante le scuole primarie. In seguito verranno esposte alcune modalità usate per verificare le competenze lessicali degli studenti e le caratteristiche delle definizioni dei bambini, spesso richieste durante tali test. Infine si farà una sintesi delle indagini riguardanti le competenze lessicali dei bambini. Lo studio di queste ricerche ha costituito, infatti, il punto di partenza per realizzare il presente lavoro. Nell'ultimo capitolo verrà presentata la ricerca sperimentale condotta su tredici classi delle scuole primarie dell'Istituto Comprensivo "Dante Alighieri" di San Canzian d'Isonzo.

1. Il lessico

La possibilità di comunicare con i propri simili, di trasmettere e decifrare messaggi linguistici è di fondamentale importanza per la vita in società: senza capire e farsi capire degli altri l'uomo non può realizzare nulla. La diversità culturale, che permea la realtà in ogni sua minima sfaccettatura, ha fatto sì che ogni comunità umana organizzi la lingua in modo diverso perché diverso è il suo modo di guardare al mondo. Ruolo chiave in questo processo è rivestito dal lessico, che non può essere un semplice elenco di etichette posizionate su concetti stabili nel tempo e nello spazio, ma è un insieme aperto, variabile e soggetto continuamente a perdite, a incrementi e a riposizionamenti: le parole perdono accezioni, ne acquistano di altre, si specializzano. Questa plasticità è conseguenza del fatto che nel tempo cambiano le necessità, "le cose da dire", e ogni lingua ha quindi «bisogno di adattare il suo patrimonio lessicografico alle [nuove] esigenze» (De Mauro 2008, p.36).

1.1. Che cos'è il lessico

1.1.1. Lessico, vocabolario, dizionario

Ogni persona, in quanto parlante di una lingua, conosce un certo numero di parole. Le parole conosciute da ogni parlante varia per quantità e qualità in base alle esperienze personali, al livello culturale, agli ambiti di interesse, e così via. Nessun parlante tuttavia, potrà mai conoscere tutte le parole della lingua, cioè nessuno potrà possedere l'intero lessico di una lingua. Se, infatti, il vocabolario è «l'insieme dei vocaboli usati da un singolo parlante [...] o da un gruppo di parlanti [...] o, in un senso ancora più restrittivo, l'insieme dei vocaboli che compaiono in un singolo testo scritto o parlato» (Casadei 2003, p. 116), il lessico è dato dalla somma e dall'intersezione di questi vocabolari o, più semplicemente, è «l'insieme delle parole di una lingua» (Ježek 2005, p.13). Adattando la distinzione saussuriana tra *langue* e *parole*, Casadei propone un'efficace proporzione in grado

di spiegare la differenza tra vocabolario e lessico, ovvero dice che «il lessico sta alla *langue*, al sistema linguistico, come il vocabolario sta alla *parole*, all'uso della lingua» (Casadei 2003, p. 115). Il lessico è infatti astratto, poiché ha dimensioni e componenti in continuo mutamento: parole di nuova formazione entrano costantemente nel lessico, sia tramite i meccanismi morfologici intrinseci alla lingua stessa, sia come prestiti provenienti da altre lingue. Per questo motivo e poiché nessuno padroneggia l'intero lessico di una lingua, determinarne le dimensioni esatte è missione quasi impossibile.

Nell'uso comune *lessico* e *vocabolario* vengono usati come sinonimi, ma, come abbiamo visto, in ambito specialistico questi due termini indicano concetti differenti. Anche il termine *dizionario* è spesso fatto coincidere con i primi due, ma anche in questo caso nelle discipline linguistiche designa un concetto ben preciso e distinto dai due precedentemente esposti. Per *dizionario*, infatti, si intende l'oggetto fisico e concreto (cartaceo o digitale) nel quale vengono raccolti i lemmi di una lingua. Il dizionario tenta di raccogliere e descrivere il lessico di una lingua, ma è evidente che ogni dizionario non può che rappresentare soltanto una parte, per quanto ampia, del lessico. «In ambito dizionaristico, l'eshaustività è da ritenersi non solo un miraggio, ma un vero abbaglio epistemologico» dice infatti Valentina Bisconti (2012, p.8). Gli autori di un dizionario sono costretti a fare delle selezioni che tengano conto dell'utenza alla quale si rivolge, della natura del dizionario, del momento storico a cui il dizionario si riferisce, o altro. Da un dizionario generale, per esempio, saranno esclusi i termini specialistici della botanica o della psicologia.

I dizionari seguono perlopiù l'ordine alfabetico per facilitarne la consultazione; tuttavia esistono anche dizionari che seguono criteri espositivi diversi, come quelli analogici (o ideologici) che «raggruppano le parole in base alla vicinanza del loro significato» (Ježek 2005, p.14). Le parole sono elencate usando la loro forma di citazione (o lemma). Il lessico, al contrario, contiene tutte le forme flesse di verbi, nomi e aggettivi, e non è organizzato alfabeticamente. Al suo interno, infatti, si possono tracciare dei raggruppamenti diversi in base al criterio preso come riferimento: quello sintattico dividerà le parole in verbi, nomi, aggettivi, ecc.;

quello semantico in aree semantiche; e quello morfologico in parole semplici e parole complesse. Dunque il lessico non può essere considerato come una semplice lista di parole (semplici, complesse e formazioni irregolari) priva di strutturazione interna.

1.1.2. Lessicologia e lessicografia

A fronte delle differenze tra lessico e dizionario si sono sviluppate due diverse discipline: la lessicologia e la lessicografia. La lessicologia è il settore della linguistica che si occupa del lessico, utilizza un approccio teorico per individuare il modo in cui il lessico è organizzato e mira dunque all'elaborazione di una teoria lessicologica, cioè di «un'ipotesi riguardo a come il lessico è strutturato, e di un modello lessicologico, cioè di un insieme di strumenti formali in grado di rappresentare questa struttura.» (Ježek 2005, p. 16). La lessicologia, per sua natura, utilizza apporti di altri settori della linguistica, come la morfologia (per quanto riguarda la struttura interna e i processi di formazione delle parole), l'etimologia (per la loro origine), la statistica linguistica (per la loro frequenza d'uso) e la semantica (per il significato e i rapporti di significato tra le parole).

La lessicografia, invece, si occupa della realizzazione dei dizionari. Deve quindi individuare il modo migliore per descrivere i significati, le proprietà grammaticali e gli usi possibili delle parole, anche quelli traslati, facendo attenzione a presentare le informazioni in modo accessibile all'utente a cui si rivolge. Nei dizionari le parole sono riportate nella loro forma di citazione, o lemma, che in italiano corrisponde per i verbi al modo infinito, per i nomi alla forma singolare (o plurale per quei nomi che hanno solo la forma plurale, come *forbici*), per gli aggettivi al maschile singolare. Il lessema è invece l'unità di base del lessico; è un'unità astratta che concretamente può assumere diverse forme: il lessema *gatto* ha sia la forma *gatto* che *gatti* (Casadei 2003, p.50).

1.1.3. La lessicalizzazione

Oltre al piano lessicale vi è quello semantico poiché la parola ha la peculiarità di avere un significato. Quest'ultimo ha una natura concettuale, che lo distingue dal referente reale al quale si riferisce: le parole, infatti, rimandano a concetti, non a oggetti reali. I rapporti che intercorrono tra significato, significante e referente sono arbitrari (non c'è nessun elemento in un albero che giustifichi l'associazione di quell'oggetto reale con la parola italiana *albero*) e l'associazione tra un concetto e una forma lessicale è detta lessicalizzazione, che ha come risultato la formazione di una parola.

Si può guardare alla lessicalizzazione sia ponendo attenzione al processo (interpretazione dinamica), sia al risultato (interpretazione statica). Guardando alla lessicalizzazione come processo è possibile individuare due accezioni possibili del termine. In senso lato, per lessicalizzazione si intende qualsiasi procedimento per cui un concetto è associato a una forma lessicale, in modo da ottenere una parola prima inesistente in quella lingua. Secondo questa interpretazione, ogni parola può essere considerata il risultato di una lessicalizzazione. In senso stretto, si considera la lessicalizzazione come il procedimento specifico per cui «una sequenza di elementi lessicali che frequentemente si presentano adiacenti in un testo, oppure una forma della flessione nominale o verbale, acquistano progressivamente valore di unità lessicale con significato autonomo, e quindi di parola.» (Ježek 2005, p. 18). In questo senso *perlopiù* e *cantante* sono delle lessicalizzazioni: *per lo più* per univerbazione è diventato l'avverbio *perlopiù*; il participio presente del verbo cantare si è lessicalizzato nel sostantivo con tratto [+ANIMATO] *cantante*. Infine, in un'interpretazione statica, *straniero* e *forestiero* sono due lessicalizzazioni che la lingua italiana ha per un unico concetto.¹

Tuttavia non esiste un'unica modalità con cui un concetto si lessicalizza in una lingua. Innanzitutto va precisato che ogni lingua 'ritaglia' la realtà in modo diverso ed è per questo che non esiste una corrispondenza biunivoca tra parole e concetti. Ogni lingua, essendo espressione di una cultura, può assegnare parole

¹ Esempio tratto da *La struttura delle parole* (2008) di Antonietta Bisetto e Sergio Scalise.

isolate a concetti che in altre lingue vengono espressi mediante perifrasi, oppure può usare un unico termine per esprimere concetti che in altre lingue vengono specificati da parole diverse. Al primo tipo appartiene il caso di *clock* e *watch*, traducibili in italiano con *orologio da parete* e *orologio da polso*; al secondo tipo è riconducibile il caso, ad esempio, dell'italiano *tempo* che in inglese può essere tradotto con *time* se ci si riferisce al tempo cronologico, con *weather* se invece si parla del tempo meteorologico.

Oltre al modo in cui segmentano i concetti, le lingue si distinguono anche per il modo in cui associano un concetto ad un termine. Vi sono, infatti, sia lessicalizzazioni sintetiche che analitiche. Per lessicalizzazioni sintetiche si intende quel processo per cui durante l'associazione concetto-parola si effettua una condensazione di più elementi in uno stesso elemento lessicale. Questo avviene in modo particolarmente evidente nei verbi: se *muovere* esprime soltanto il MOTO, *andare* specifica che il movimento ha anche una DIREZIONE, così come *correre* oltre al MOTO specifica anche la MANIERA. In questo modo più informazioni vengono compresse in un'unica forma lessicale. Al contrario, con lessicalizzazione analitica ci si riferisce a quel processo per cui un unico concetto viene veicolato usando più forme lessicali. Un esempio di questa modalità sono le espressioni *avere paura* e *essere al sicuro*. Tuttavia, spesso le lingue usano entrambi i procedimenti; in italiano, ad esempio, si può dire sia *uscire* che *andare fuori*, veicolando lo stesso concetto rispettivamente con una forma sintetica e una analitica.

Si può inoltre stabilire una distinzione sulla base della relazione semantica esistente tra il referente reale (designato) e la parola. In questo modo si individuano lessicalizzazioni descrittive e lessicalizzazioni etichettanti. Nelle prime la parola descrive in qualche modo il designato, nelle seconde la parola indica l'oggetto per mezzo di un'etichetta. La lessicalizzazione descrittiva è più frequente nelle lingue del Nord e Centro America, ma si possono fare anche degli esempi tratti dalla lingua italiana, come *patente di guida* e *lavoratore* (il suffisso -tore indica la persona che svolge l'azione indicata dal verbo di base). *Medico* e *finestra* sono invece lessicalizzazioni etichettanti, poiché non c'è nessun elemento che rimandi direttamente al designato.

1.2. Che cosa sono le parole

Fin qui abbiamo tentato di delineare che cosa i linguisti intendono quando parlano di lessico e lo abbiamo fatto parlando genericamente di 'parole', ma che cosa intendiamo con 'parola' non è di facile definizione, anche se intuitivamente tutti sanno di che cosa si tratta.

Nel cercare una definizione di *parola*, sono stati utilizzati criteri differenti. Applicando un criterio ortografico, la parola può essere intesa come «ciò che è compreso tra due spazi bianchi» (Bisetto e Scalise 2008, p. 58), ma appare subito evidente il limite di questa definizione: può essere applicata solo alle lingue dotate di scrittura. Anche definire la parola come elemento dotato di 'significato unitario' non è soddisfacente poiché da questa vaga definizione vengono escluse le parole grammaticali, che da sole non possono formare un enunciato. La definizione basata sul criterio semantico si trova in difficoltà anche nella spiegazione dei casi di omonimia (cioè quando due parole condividono la forma fonologica e/o grafica ma hanno significato diverso) e di polisemia (cioè quando un'unica forma ha più significati). Da un punto di vista fonologico, inoltre, le parole si caratterizzano per il fatto di avere un accento primario. Questo, se in alcune lingue, come il francese, ha una posizione fissa e può quindi costituire un criterio valido per distinguere una parola dall'altra, in altre ha posizione libera (come in italiano), non consentendo di fare il medesimo ragionamento. Un altro criterio fonologico valido per alcune lingue e non generalizzabile, si basa sul fatto che certe sequenze di suoni si possono trovare solo in alcune posizioni all'interno della parola, ad esempio «in sanscrito la sequenza "vocale-vocale" è possibile solo in posizione iniziale di parola» (Bisetto e Scalise 2008, p.59).

La parola è dunque un'entità complessa che sta al confine tra fonologia, semantica e sintassi. La problematicità di questa nozione è evidente soprattutto se si considerano espressioni come *ferro da stiro* o *tirare le cuoia* che pur essendo composte graficamente da più parole, da un punto di vista semantico costituiscono un'unità. Per cercare di spiegare questo tipo di fenomeni, recentemente la nozione di parola è stata interpretata in modo prototipico,

ritenendo che «vi sia un *continuum* [...] lungo il quale si collocano vari tipi di forme complesse che saranno ritenute tanto più parole quanto più presentano le caratteristiche peculiari della parola prototipica» (Casadei 2003, p.125).

Partendo dal presupposto che la parola da un lato costituisca un'unità atomica, dall'altro sia composta da parti più piccole, come le sillabe, i morfemi e i fonemi, sono state individuate due proprietà principali che ogni parola, per essere considerata tale, deve avere: la coesione tra le parti costituenti e l'ordine fisso tra le parti costituenti. Detto diversamente, non si possono spostare i costituenti della parola (tana/*taan) né elementi esterni possono essere inseriti all'interno della parola (tana/*tanza). Anche un'altra proprietà viene spesso considerata peculiare della parola: l'autonomia, la capacità cioè di veicolare un senso compiuto e di comparire da sola in un testo. In questo la parola si distinguerebbe dai morfemi legati.

In conclusione dunque, una definizione soddisfacente in grado di comprendere tutti questi aspetti riguardanti le parole non è ancora stata scritta, rendendo tutt'ora questo concetto banale e ovvio a un primo sguardo, sfuggibile e complesso ad un'analisi più approfondita.

1.2.1. Tipi di parole

Già durante l'Antica Grecia i filosofi si sono adoperati per individuare delle classi di parole e proprio dalla tradizione classica arriva fino ai giorni nostri la distinzione delle parti del discorso. Per l'italiano queste sarebbero tradizionalmente nove: nome, verbo, aggettivo, pronome, articolo, avverbio, preposizione, congiunzione e interiezione. Questa distinzione mira a individuare le parti del discorso declinabili o coniugabili (le prime cinque dell'elenco) e quelle invariabili (le ultime quattro). Tradizionalmente, inoltre, i nomi designano entità o sostanze, i verbi azioni o processi e gli aggettivi proprietà e qualità.

Questa categorizzazione basata su un criterio formale non è tuttavia l'unica possibile. Considerando il loro significato, si possono individuare due grandi gruppi di parole: le parole contenuto e le parole funzione. Le prime hanno un contenuto

lessicale vero e proprio, mentre le seconde chiariscono le relazioni grammaticali. Per questo sono anche rispettivamente chiamate parole piene e parole vuote. Questa nomenclatura però non evidenzia adeguatamente il ruolo fondamentale che le parole funzione svolgono all'interno della frase. Senza il loro contributo, infatti, il significato della frase diventa opaco oppure ambiguo. Dunque le parole contenuto hanno un significato lessicale (sono semanticamente autonome), quelle funzione un significato grammaticale (acquistano significato in relazione alle parole contenuto tra le quali compaiono). Vi sono comunque casi limite, come per esempio gli ausiliari, che pur essendo verbi, non veicolano un significato lessicale, ma solo grammaticale. Anche tra avverbi e preposizioni ci sono casi particolari: gli avverbi in -mente sono più vicini alle parole contenuto rispetto ad altri avverbi, come ad esempio *non*.

Questa distinzione è strettamente collegata con un'altra: quella tra classe aperta e classe chiusa. Per esempio nomi, verbi e aggettivi appartengono alla classe aperta in quanto ammettono costantemente l'entrata e l'uscita di elementi; al contrario le parole funzione, che costituiscono la classe chiusa, sono numericamente finite. Anche l'introduzione di nuove parole funzione in una lingua è possibile, ma è un fenomeno più raro e soprattutto più lento, che di solito avviene nel corso di secoli. Dal punto di vista genetico, invece, si possono distinguere tre tipi di parole: quelle che si sono costituite con le regole di formazione di parola specifiche della lingua (sono soprattutto le regole morfologiche di derivazione e composizione); quelle nate dalla progressiva fissazione dei rapporti tra due parole semplici che cooccorrono frequentemente in un testo (*fino a che* > *finché*); quelle importate da altre lingue, i cosiddetti prestiti (molto diffusi attualmente in italiano sono gli anglicismi, come *shopping*).

Se poi si guarda alle parole da un punto di vista morfologico si può distinguere tra parole semplici e parole complesse. Le parole semplici sono quelle composte da un unico morfema lessicale libero (es. *ieri*) o da un morfema lessicale legato e un morfema flessivo (es. *casa*), quindi non sono il prodotto né di una derivazione né di una composizione. Le parole complesse, invece, sono derivate o composte e perciò sono «costituite da un morfema lessicale e da almeno un altro morfema

lessicale e/o derivazionale, oltre a eventuali morfi flessivi» (Ježek 2005, p.40). Esempi di parole derivate sono: *collaudatore, fiorista, spazioso*. Sono composti invece parole come *capotreno, sottoscala, taglialegna*.

Un caso di confine tra morfologia e sintassi è rappresentato dalle polirematiche (o parole sintagmatiche). Sono lessemi complessi, formati da più parole e il cui significato non è ricavabile in modo compositivo dai componenti. Sono un normale meccanismo di estensione del lessico, come derivazione e composizione, e sono un fenomeno quantitativamente rilevante: il Gradi² individua 63.000 polirematiche italiane, escludendo dal conteggio quelle dei vocabolari specialistici (le polirematiche si trovano infatti in tutti i registri linguistici, da quello colloquiale a quello specialistico) e quelle importate da altre lingue (come *big bang* o *pro capite*). Esistono polirematiche appartenenti ad ogni categoria lessicale: nomi (*macchina da scrivere, sala da pranzo*), verbi (*fare acqua, dare retta*), aggettivi (*acqua e sapone, campato in aria*), avverbi (*mano a mano*), congiunzioni (*dal momento che*), preposizioni (*a causa di*).³ Le polirematiche, proprio come le parole isolate, non ammettono inserzioni tra i costituenti (se non 'leggeri', come gli avverbi semplici, ad es. *subito*), né permettono di scambiare l'ordine dei costituenti (*rotonda tavola) (cfr. § 1.2.).

Vi sono infine le cosiddette semiparole che vengono definite come «unità del lessico che sono forme legate ma che hanno tuttavia un valore lessicale pieno, non di tipo grammaticale» (Bisetto e Scalise 2008, p.63). Le semiparole sono di origine latina o greca e tutt'oggi sono molto produttive; si tratta per esempio di *antropo-* (di origine greca) di *antropologo* o *-ficio* (di origine latina) di *caseificio*. Si differenziano dagli affissi in quanto, pur essendo forme legate che non possono comparire da sole all'interno della frase, hanno un significato lessicale, e non solamente relazionale. Le parole composte formate con questi elementi sono dette composizioni neoclassiche.

² Grande dizionario italiano dell'uso.

³ Gli esempi sono tratti perlopiù da Casadei (2003).

1.3. La stratificazione del lessico

Torniamo ora a considerare il lessico nel suo insieme. Questo non è una realtà uniforme ed omogenea, bensì un insieme eterogeneo all'interno del quale è possibile individuare strati diversi sia in rapporto all'origine delle parole, sia considerando la frequenza d'uso.

1.3.1. Stratificazione per origine

Usando il criterio etimologico, si possono individuare all'interno del lessico vocaboli di diversa origine e formazione. Per l'italiano si possono delineare tre grandi raggruppamenti di parole: quelle di derivazione latina (diretta e indiretta), quelle prese da altre lingue e quelle nate dalle regole di espansione del lessico proprie dell'italiano.

La maggior parte delle parole italiane è arrivata dal latino con due modalità: da un lato per tradizione diretta e ininterrotta, dall'altro per tradizione indiretta o dotta. Nel primo caso l'uso vivo del latino nel corso dei secoli ha prodotto cambiamenti di forma tali da dare origine alle parole romanze e, in particolare nel nostro caso, alle parole italiane. Sono quindi cambiamenti nati spontaneamente tra le conversazioni delle persone. Il secondo è il caso delle parole entrate 'forzatamente' in italiano in un momento storico in cui il latino non era più parlato ma letto, e in alcuni casi scritto, solamente dagli intellettuali e dalle persone colte. Queste parole sono dunque formalmente più vicine al loro corrispettivo latino. Esistono anche casi in cui la stessa parola latina ha dato origine a due parole, talvolta anche con significati diversi, una per via popolare e una per via dotta, sono i cosiddetti allotropi (es. da *oculum* deriva sia *occhio* per via diretta che *oculare* per via dotta).

I prestiti o forestierismi sono invece le parole prese da altre lingue. I prestiti possono adattarsi alla forma fonologica della lingua d'arrivo (prestiti adattati) oppure mantenere la loro forma originaria (prestiti non adattati), ma questo dipende soprattutto dall'epoca del prestito. *Caffè* ha una forma completamente

assimilata al resto del lessico italiano, tanto che non viene più nemmeno percepito come prestito. Al contrario *weekend* mantiene la sua veste fonologica inglese. Le lingue da cui sono tratti i prestiti italiani sono le più diverse: francese, spagnolo, tedesco, arabo, turco, russo, giapponese. Ma le due lingue più rappresentate nel lessico italiano sono il greco antico e l'inglese. Dal primo deriva la maggior parte della terminologia medica e molti vocaboli scientifici e letterari, dal secondo le parole della tecnologia e non solo. Tuttavia non sarebbe sempre necessario ricorrere a parole straniere per esprimere certi concetti. Se, infatti, esistono prestiti di necessità che introducono concetti e oggetti nuovi o di altre culture, vi sono anche prestiti di lusso che non introducono concetti nuovi, ma vengono accolti solo in virtù di un percepito prestigio della lingua di provenienza. Oltre ai prestiti, esistono anche i calchi. In queste formazioni la struttura della parola straniera viene copiata usando gli elementi lessicali già presenti in italiano. L'esempio tipico è rappresentato da *grattacielo* che strutturalmente corrisponde perfettamente all'inglese *skyscraper*.

Infine vi sono le formazioni endogene, le parole create con i meccanismi di formazione di parola propri della lingua: tramite i processi di derivazione (es. *giornalaio*) e di composizione (es. *apribottiglie*). Tra queste trovano posto anche quelle formazioni, presenti soprattutto nel vocabolario tecnico-scientifico, che si servono delle semiparole (cfr. § 1.2.1.). È il caso ad esempio di *fotografia*, *cardiochirurgo*, *multifunzione*.

1.3.2. Stratificazione per frequenza d'uso

Come abbiamo già detto, nessuno conosce l'intero lessico di una lingua. Ma quali sono le parole più conosciute e più usate dai parlanti? La risposta a questo quesito viene dalla statistica linguistica, la disciplina che si occupa di individuare nel lessico gli strati di parole più o meno conosciute applicando metodi statistici. Preso un campione di testi scritti e/o orali, viene calcolata la frequenza con cui ciascuna parola compare nel campione, la sua dispersione (cioè quante volte compare nei singoli testi) e infine il fattore d'uso che corrisponde al prodotto di

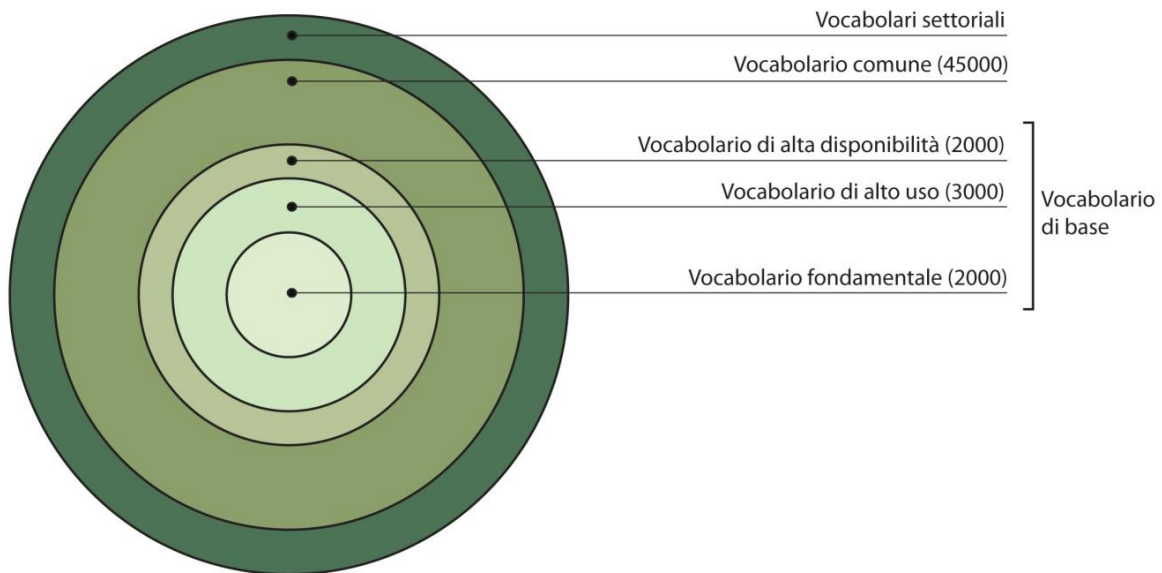
frequenza e dispersione. In questo modo le parole più frequenti saranno quelle più frequentemente usate nel maggior numero di testi diversi.

Considerando quindi l'uso, il lessico può essere rappresentato come una serie di cerchi concentrici. Al centro troviamo i circa 2000 vocaboli di massima frequenza che costituiscono il lessico fondamentale. Queste parole compaiono nel 90% dei discorsi e dei testi scritti e oltre a nomi, verbi e aggettivi molto frequenti, comprendono le parole funzione e gli ausiliari. Qualunque italiano con licenza elementare conosce il vocabolario fondamentale. Subito dopo troviamo il vocabolario di alto uso che contiene altri 3000 vocaboli molto frequenti. A seguire è stato individuato il vocabolario ad alta disponibilità, che contiene parole usate non molto frequentemente, ma sicuramente presenti nel lessico mentale dei parlanti. Qui sono contenuti 2000 vocaboli legati alla vita quotidiana, poco usati nello scritto perché considerati troppo banali e poco usati nel parlato perché spesso sostituiti da cenni, gesti o dimostrativi, come ad esempio *forchetta* e *ciabatta*. Vocabolario fondamentale, d'alto uso e di alta disponibilità costituiscono quello che viene chiamato il vocabolario di base, formato da circa 7000 parole conosciute da chiunque abbia completato l'istruzione di base (fino alla scuola secondaria di primo grado) e utilizzate nel 98% dei discorsi. Del vocabolario di base si parlerà ancora e più dettagliatamente al paragrafo 1.7.

Lo strato successivo è formato dai 45.000 vocaboli appartenenti al vocabolario comune. Qui troviamo le parole note alla maggior parte di coloro che hanno un'istruzione medio-superiore, indipendentemente dalla professione che svolgono o dalla loro provenienza geografica. A questo strato appartengono anche parole provenienti dai linguaggi tecnico-scientifici ma ormai diffuse anche tra i non specialisti. Qualche esempio di parola del vocabolario comune: *galleggiante*, *germogliare*, *ecologico*, *insulina*, *radiografia*.

Vi è poi un ampio strato formato dai vocabolari settoriali, come quelli tecnico-scientifici e letterari. Ogni disciplina conta infatti milioni di parole, conosciute quasi esclusivamente dagli specialisti. Ristretti ad ambiti particolari sono inoltre i regionalismi e i vocaboli dialettali entrati nel lessico italiano.

Gli starti più esterni del lessico sono infine composti dalle parole a basso uso e dai vocaboli obsoleti, cioè quelli che pur non essendo più usati vengono comunque registrati nei dizionari.



1.4. Significato, semantica e campo semantico

Finora ci siamo concentrati soprattutto sugli aspetti propriamente lessicali, ma l'altro versante delle parole riguarda il loro significato. Specificatamente di questo si occupa la semantica, il settore della linguistica dedicato allo studio del significato.

Il concetto di significato, a guardare in profondità, è controverso poiché si trova al confine tra linguaggio, pensiero e realtà. In generale tuttavia si può dire che il significato è l'informazione trasmessa da una parola o da un'espressione linguistica. Diversamente che nel linguaggio corrente, in linguistica *sensu* e *significato* sono due concetti distinti. Se il *significato* è «il contenuto che un segno linguistico ha nella *langue*», nel sistema linguistico, il *sensu* è «il contenuto che il segno veicola in una specifico atto di *parole*, cioè [...] in una particolare situazione comunicativa» (Casadei 2003, p. 34).

Nella maggior parte dei casi, tuttavia, le parole non hanno un solo significato, ma possono essere usate in accezioni diverse. La polisemia (cioè la caratteristica di una parola di avere più accezioni di significato) è infatti un fenomeno diffuso in tutte le lingue.

I tipi di significato che una parola può assumere sono diversi. Una prima distinzione possibile è tra significato denotativo e significato connotativo. Il significato denotativo è quello oggettivamente espresso o descritto dalla parola o dall'espressione linguistica, è la parte di significato «condivisa dai membri di una comunità linguistica e che consente quindi di fatto lo scambio di informazione che i parlanti attuano attraverso l'uso di questa parola» (Ježek 2005, p. 48). Il significato connotativo, invece, aggiunge contenuti al significato di base, può far trasparire ad esempio l'atteggiamento del parlante nei confronti del referente (*padre* e *papà* condividono il significato denotativo ma hanno diverso significato connotativo) o può dare informazioni riguardo lo stile della comunicazione (*cavallo* verrà usato in contesti quotidiani, *destriero* in ambiti letterari).

Un'altra distinzione possibile, come già esposto precedentemente (cfr. § 1.2.1.), è quella tra significato lessicale e grammaticale.

Infine una parola può assumere in una frase un significato letterale, che coincide con quello denotativo, oppure uno non letterale. In questo secondo caso la parola viene caricata di riferimenti più ampi rispetto a quelli che le sono propri, soprattutto per mezzo di figure retoriche come la metafora, la metonimia o l'ironia.

Facendo riferimento al loro significato, le parole possono essere accomunate dal fatto di 'coprire' una certa area concettuale. Tale raggruppamento è chiamato campo semantico (o anche area semantica). Per esempio *spada*, *scudo*, ed *elmo* fanno parte del campo semantico della guerra.

1.5. Il lessico mentale

Scopo della linguistica è sì capire il funzionamento delle lingue, ma è anche quello di riuscire a spiegare i processi mentali che sottostanno alla produzione e alla

comprensione linguistica. Finora abbiamo guardato al lessico da un punto di vista strettamente linguistico e infatti lo abbiamo definito come l'insieme delle parole di una lingua. «L'insieme delle parole conosciute da un parlante» (Bisetto e Scalise 2008, p.17) costituisce invece il lessico mentale del parlante, nel quale rientrano anche l'insieme delle informazioni sulle parole che servono ad utilizzare correttamente le parole stesse.

Dagli studi condotti in ambito psicolinguistico e neurolinguistico è emersa l'esistenza più che di un unico lessico mentale, di quattro lessici: due in entrata (fonologico per le parole pronunciate a voce alta e ortografico per le parole presentate in forma scritta) e due in uscita (anche in questo caso uno fonologico e uno ortografico). Questi sarebbero poi connessi con un sistema semantico-concettuale in grado di assegnare alla parola il suo significato (Denes 2009). Le informazioni grammaticali di ogni parola (genere, ruolo tematico, ecc.) sarebbero inoltre memorizzate separatamente dalla forma fonologica.

Ci si è anche chiesti se le parole siano memorizzate nel lessico in forma unitaria o decomposta (radici da una parte e morfemi grammaticali dell'altra). Attualmente la maggioranza degli studi è propensa ad assumere una posizione intermedia: le parole più frequenti e quelle irregolari sarebbero memorizzate in forma unitaria, mentre quelle poco frequenti e le regolari in forma scomposta; per recuperare queste ultime nella memoria, verrebbero applicate quindi delle regole in grado di assegnare alla radice il corrispondente affisso. È comunque ormai appurato che caratteristiche semantiche e di struttura interna della parola influenzano sia il modo in cui questa è memorizzata nel lessico mentale sia i processi di riconoscimento, comprensione e produzione.

Le ricerche in questo ambito sono svolte soprattutto su pazienti con disturbi del linguaggio poiché «gli errori, sia di produzione che di comprensione di parole, si possono considerare [...] delle fessure, attraverso le quali è possibile intravedere [...] i processi cognitivi sottostanti il linguaggio» (Denes 2009, p.87)

1.6. I dizionari di frequenza della lingua italiana e i dizionari di nuova generazione

A cominciare dagli anni Settanta del Novecento si è diffuso anche in ambito italiano il filone di ricerca che mirava a realizzare dizionari in modo completamente nuovo rispetto al passato: i dizionari di frequenza. In precedenza c'erano già stati alcuni contributi, ma erano stati effettuati tramite spoglio manuale delle fonti. Da ricordare, ad esempio, quelli di Knease (1933) e di Migliorini (1943). Il primo è uno spoglio di 4000 parole tratte da fonti letterarie, il secondo una selezione delle 1500 parole considerate, dal gusto linguistico dell'importante storico della lingua, come fondamentali nella lingua italiana dell'epoca (soprattutto in prospettiva didattica).

Con il diffondersi anche tra gli umanisti delle tecniche di elaborazione automatica, si vengono dunque a creare i presupposti tecnici e culturali per realizzare veri e propri dizionari di frequenza della lingua italiana e nel 1972 viene pubblicato il *Lessico di frequenza della lingua italiana contemporanea* (LIF) curato da Bortolini, Tagliavini e Zampolli. Quest'opera segna un salto di qualità rispetto alle precedenti: è costituito dallo spoglio di 500.000 parole in forma scritta della lingua italiana contemporanea, dal quale sono stati tratti circa 5000 lemmi risultati più frequenti. Le fonti si distinguono per il fatto di riferirsi all'attualità (1947-1968) e sono suddivise in cinque settori: romanzi, periodici, sussidiari per la scuola elementare, testi teatrali e sceneggiature cinematografiche. Questi ultimi due tipi di testi erano all'epoca ciò che più si avvicinava all'italiano parlato, pur trattandosi di una lingua simulata. Da ogni categoria sono state tratte ugual numero di occorrenze (100.000).

Due anni dopo Juilland e Traversa pubblicano il *Frequency Dictionary of Italian Words* (FDI), che però si riferisce al periodo 1920-1944. Nel 1977 inoltre viene pubblicato il *Vocabolario Fondamentale della lingua italiana* di Sciarone che ai *corpora* di LIF e FDI aggiunge un nuovo *corpus*, riferito agli anni 1973-1974, di altre 500.000 occorrenze. Questo *Vocabolario* presenta una lista di frequenza di 2500 parole.

Sempre dedicato all'italiano scritto è il *Vocabolario Elettronico della Lingua Italiana* (VELI) a cura della Direzione Ricerca Scientifica e Tecnologica IBM Italia edito nel 1989. Quest'opera «presenta un *corpus* con un'estensione vastissima, circa 26 milioni di occorrenze, ma con un orientamento molto forte verso lo scritto formale, data la tipologia testuale prescelta, costituita fundamentalmente da settimanali informativi, settimanali economici, notiziari ANSA e simili» (D'Agostino 1998, p. 12). La lista così ricavata è di 10.000 lemmi.

Il primo lessico dedicato alla lingua parlata arriva nel 1993 ed è il *Lessico di Frequenza dell'Italiano Parlato* (LIP) di De Mauro, Mancini, Vedovelli e Voghera. Si presenta come «strumento volto alla verifica sperimentale delle ipotesi relative ai processi di unificazione linguistica dell'Italia, con particolare attenzione, alla verifica della diffusione dell'italiano nel suo registro parlato» (D'Agostino 1998, p. 13). Per realizzare il LIP sono state prese in considerazione cinque diverse tipologie di testi orali raccolti in quattro città italiane (Milano, Roma, Firenze e Napoli) in modo da rappresentare realtà geografiche, demografiche e di prestigio linguistico diverso. Anche in questo caso il corpus è costituito da 500.000 parole.

A partire dalla seconda metà degli anni Novanta si registra una nuova tappa del percorso innovativo in ambito lessicografico: fanno, infatti, la loro comparsa i dizionari di nuova generazione, così chiamati perché contengono informazioni prima tralasciate dai dizionari. Per ciascun lemma vengono fornite, oltre al significato, indicazioni sugli usi sintattici e pragmatici. In questo modo (Ferreri 2006, p. 132)

i dizionari vengono [...] a configurarsi sempre più non solo come rappresentazione del significato lessicale ma come spaccato della lingua nel suo complesso da cui è possibile cogliere le forze agenti nel determinarsi dei sensi e del loro consolidarsi in accezioni stabili, la pressione degli usi e delle consuetudini di una comunità.

Applicando ricerche di statistica linguistica, infatti, questi nuovi strumenti lessicografici permettono di osservare anche la stratificazione degli usi delle parole, cogliendo i processi in atto sul piano pragmatico. Per rendere ciò si è ricorsi alle marche d'uso, che vengono assegnate seguendo criteri di frequenza tratti appunto dalla statistica linguistica. Ne deriva una stratificazione degli usi

dell'intera comunità, che non hanno a che fare con la percezione che il singolo parlante ha.

Il *Grande Dizionario Italiano dell'Uso* (Gradit) di Tullio De Mauro edito dalla Utet nel 1999 è forse l'esempio più noto tra i dizionari di nuova generazione. In esso ogni lemma è accompagnato da una marca d'uso: FO vocabolario fondamentale, AU di alto uso, AD di alta disponibilità, CO comune, TS tecnico specialistiche, ES esotico, DI dialettale, RE regionale, LE di uso letterario, BU di basso uso, OB obsoleto. In questo modo le parole accomunate dalla stessa marca identificano uno strato del lessico (cfr. § 1.3.2.) e rendono evidenti quali parole vengono pronunciate, scritte e capite dai parlanti più frequentemente. Il rapporto fondamentale, da sempre esistente, tra lingua e comunità di parlanti viene così per la prima volta evidenziato anche nel dizionario, permettendo di «configurare il lessico di una lingua come il regno delle relazioni delle parole tra loro e delle parole con i parlanti e le loro vicende storiche e culturali» (Ferrerri 2006, p. 132).

1.7. Il vocabolario di base

Il vocabolario di base della lingua italiana, che d'ora in poi citeremo come VdB per comodità, «è l'insieme delle parole che si ritiene costituiscano il nucleo essenziale della lingua» (Thornton *et al.* 1997, p.27). Si tratta di uno strumento lessicografico accomunabile per certi aspetti ai dizionari (riporta soltanto i lemmi e non le forme flesse), per altri ai lessici di frequenza (include i lemmi più frequenti dell'italiano). Presenta tuttavia dei limiti: non considera le forme flesse e non specifica quale accezione o quali accezioni del lemma appartengono al VdB, generando una certa ambiguità semantica di parole che, proprio in virtù della loro frequenza d'uso, sono spesso polisemiche, il corpus di riferimento è un po' datato (decenni centrali del Novecento), non tutti i lemmi sono stati selezionati seguendo gli stessi criteri. Nel Gradit (1999) invece De Mauro assegna la marca d'uso ad ogni singola accezione del lemma.

1.7.1. Come è stato realizzato

La base di partenza da cui sono state tratte le indicazioni di frequenza delle parole è il LIF (cfr. § 1.6.), poiché all'epoca in cui è nata l'idea di realizzare il VdB era il lessico di frequenza che meglio rappresentava l'italiano moderno effettivamente usato dai parlanti. Come abbiamo già ricordato, il LIF è stato realizzato partendo da un *corpus* di 500.000 occorrenze, selezionate in ugual misura da cinque tipologie testuali diverse, tutte pubblicate tra il 1947 e il 1968. Le 500.000 occorrenze sono state lemmatizzate: le forme flesse sono state ricondotte quindi alla loro forma di citazione, che per i nomi corrisponde alla forma singolare, per gli aggettivi al maschile singolare, per i verbi all'infinito. Per ogni forma è stato calcolato l'indice di frequenza e, sommando le frequenze di tutte le forme riconducibili allo stesso lemma, si è ottenuta la frequenza complessiva del lemma. Per ogni forma e per ogni lemma è stato inoltre calcolato l'indice di dispersione, che mostra quanto una parola compaia nei diversi sottocorpora. Una parola può comparire molte volte in un testo e mai in un altro in base all'argomento trattato e con l'indice di dispersione si rende conto di questo fenomeno, evitando di distorcere l'uso reale di quella parola. L'indice di dispersione, infatti, assume valori compresi tra 0, quando le occorrenze di un lemma compaiono in una sola tipologia testuale, e 1, quando al contrario le forme di una parola ricorrono in tutti e cinque i sottocorpora. Infine è stato calcolato l'indice d'uso, dato dal prodotto di frequenza e dispersione, che assume valori più bassi nelle parole che non sono uniformemente distribuite nei sottocorpora e che rende la reale 'importanza' di una parola. A questo punto sono stati selezionati i 5356 lemmi più frequenti (superavano un certo valore nell'indice d'uso).

Il lemmario realizzato secondo questi criteri dagli autori del LIF è stato utilizzato come riferimento da De Mauro e i suoi collaboratori per realizzare il VdB. Hanno considerato le parole con indice d'uso superiore o pari a 2 (5001 lemmi) e le hanno divise in due fasce, la prima di 2000 vocaboli e la seconda con i restanti 3000. Stefano Gensini ed Emilia Passaponti, hanno collaborato con De Mauro occupandosi di verificare che queste parole fossero realmente comprese da chi

aveva non più e non meno che la licenza media (studenti e adulti), ipotesi da cui partiva il progetto di De Mauro. Le prime 2000 parole sono risultate conosciute e correttamente usate da tutti e hanno dunque formato il vocabolario fondamentale. Delle 3000 parole della seconda lista di frequenza, molte sono risultate non comprese da metà dei soggetti esaminati e sono state perciò eliminate, riducendo il vocabolario d'alto uso a circa 2900 vocaboli. Nel VdB compaiono però anche parole che o non sono presenti nel LIF o se lo sono hanno rango superiore a 5001: «sono le parole che diciamo o scriviamo raramente, ma che pensiamo con grande frequenza» (De Mauro 2003, p. 162) e che per questo rischiano di non venir registrate nelle liste di frequenza. Per individuarle Gensini e Passaponti hanno letto insieme ad alta voce tutti i lemmi di un dizionario d'italiano comune compresi tra una parola fondamentale e l'altra e per ogni lemma hanno segnato autonomamente se a loro giudizio era familiare o meno. Dal confronto e dalla discussione delle due liste è stato realizzato un nuovo elenco sottoposto al giudizio di De Mauro che si è preoccupato di eliminare le parole legate alla sola realtà toscana. La nuova lista ottenuta dopo questi passaggi è stata sottoposta al giudizio di coloro che erano in possesso della licenza media. Alla fine, la lista di circa 1750 parole così selezionate è andata a formare il vocabolario di alta disponibilità.

Il VdB è stato pubblicato per la prima volta in appendice alla prima edizione di *Guida all'uso delle parole* (1980) di Tullio De Mauro, ma il lavoro di verifica della sua comprensibilità è continuato ancora presso le scuole medie. Mentre le parole fondamentali sono rimaste invariate, quelle ad alto uso sono state ridotte a 2750. Anche il vocabolario di alta disponibilità è stato modificato, aggiungendo sia parole familiari che erano sfuggite (ad es. *peperone*) sia altre del lessico aritmetico, medico e politico molto diffuse.

Con questo strumento non si intendeva costruire una gabbia entro la quale intrappolare ogni discorso, ma fornire un riferimento al quale affidarsi durante la stesura di testi e discorsi: usando le parole del lessico fondamentale verremo capiti da chiunque abbia completato almeno la scuola elementare, usando quelle dell'intero vocabolario di base verremo capiti molto probabilmente da chi ha

almeno la licenza media. Questo perché la scelta delle parole gioca un ruolo decisivo nella possibilità di capire un testo, e se «non pare esserci limite noto alla possibilità di oscurare un discorso e di rubare agli altri la possibilità di capirlo» (De Mauro 2003, p. 163), non dobbiamo dimenticare che obiettivo principale della comunicazione dovrebbe essere quello di far passare un messaggio, di farsi capire.

1.7.2. Dati quantitativi del VdB

Nel 1994 Anna Thornton, Claudio Iacobini e Cristina Burani hanno realizzato una base di dati per rispondere a una nascente necessità di disporre di dati quantitativi sulla composizione lessicale dell'italiano. L'opera, intitolata *BDVDB. Una base di dati sul vocabolario di base della lingua italiana* e riedita nel 1997, ha in allegato un *floppy disk* che contiene l'intera base di dati realizzata basandosi sul VdB. Nel floppy disk i lemmi del VdB sono presentati corredati da molte informazioni: la lunghezza del lemma in grafemi, il numero di sillabe, la posizione dell'accento, la struttura morfologica (parola semplice, prefissata, suffissata, composta, ecc.), la categoria grammaticale, la classe flessiva, il genere dei nomi, la lingua d'origine, la trasmissione dotta o popolare, la data di prima attestazione, il secolo di prima attestazione, la fascia di frequenza alla quale appartiene il lemma nel VdB, la disambiguazione semantica di alcuni lemmi, l'inversione della stringa di lettere. In questo modo, nella descrizione della lingua, si possono fare considerazioni sia quantitative che sulle unità sublessicali (numero di sillabe più frequente, posizione dell'accento, desinenze, ecc.).

Di seguito esporremo alcune considerazioni quantitative che Thornton e collaboratori hanno riportato nel volume come esempio delle potenziali ricerche che si possono fare a partire dalla loro base di dati.

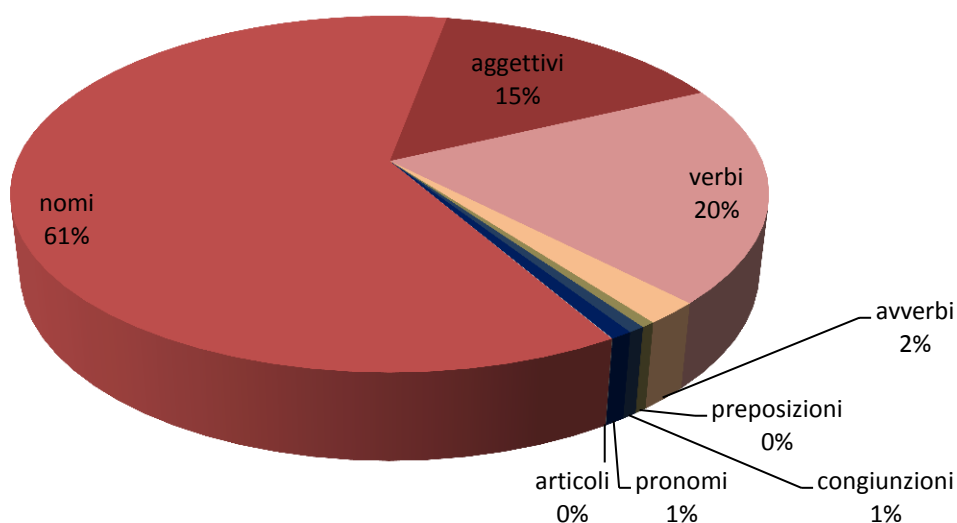
Innanzitutto hanno calcolato la distribuzione nelle diverse categorie grammaticali delle parole del VdB, prendendo in considerazione anche le fasce d'uso.

Cat.grammaticale	FO (%)	AU (%)	AD (%)	VdB (%)
Nome	46,7	58,6	75,4	60,7
Aggettivo	14,9	18,2	11,3	14,9
Verbo	25,8	20,7	12,8	19,6
Avverbio	5,4	1,0	0,2	2,0
Preposizione	1,3	0,2	0,1	0,5
Congiunzione	1,5	0,3	0,04	0,6
Pronome	2,5	0,03	0,04	0,8
Articolo	0,1	-	-	0,02

Tabella 1: Distribuzione percentuale dei lemmi del VdB nelle categorie grammaticali. Dati da Thornton *et al.* (1997).

Come si vede nella tabella sopra riportata, i nomi sono la categoria grammaticale più rappresentata sia nel VdB nel suo complesso che nelle diverse fasce d'uso, seguono verbi e aggettivi; tuttavia in termini percentuali le differenze tra nome e verbo variano nelle diverse fasce: i nomi si concentrano soprattutto nel vocabolario di alta disponibilità, mentre la maggior parte dei verbi appartiene al vocabolario fondamentale. Le categorie grammaticali minori, costituite perlopiù da parti invariabili del discorso, si concentrano soprattutto nel vocabolario fondamentale.

Categorie grammaticali del VdB



Da un punto di vista morfologico il 36,6% delle parole del VdB sono semplici, mentre le restanti sono complesse, in particolare: il 27,2% sono lemmi suffissati, l'11,9% prefissati, il 5,2% composti e il 19,1% a suffissazione zero. In generale quindi tra le parole complesse quelle suffissate prevalgono. Considerando però le classi grammaticali separatamente, si nota che nei verbi prevale la suffissazione.

Se si guarda invece alla formazione etimologica delle parole, il VdB risulta costituito per il 52,2% da parole di origine latina, per il 34,3% da parole formate direttamente in italiano e per il restante 13,5% da parole provenienti da: altre lingue romanze, dialetti italiani, lingue germaniche, greco, lingue non indoeuropee e fonti sconosciute o incerte.

Infine, considerando la costituzione del VdB nel corso dei secoli, la maggior parte delle parole di cui è formato sono entrate in uso tra i secoli XIII e XIV (l'80% del vocabolario fondamentale risale a questi due secoli), confermando ulteriormente il valore speciale che questi secoli rivestono nella storia della lingua italiana. Seguono i secoli XVI e XIX per il numero di nuove immissioni nel lessico italiano. Le parole del VdB sono dunque perlopiù parole con una storia lunga secoli, ma non mancano, come è lecito aspettarsi, numerosi esempi di vocaboli entrati più di recente. Sono soprattutto parole che rendono conto delle nuove esperienze quotidiane, come *autobus*, *barista* e *accendino*. Intrecciando invece le informazioni relative ai secoli di immissione con quelle dell'origine, si può evidenziare come tra il XV e il XVI secolo si verifichi una svolta: le parole non sono più attinte principalmente dal latino, ma prevalgono le formazioni endogene dell'italiano. Da questo momento in poi il latino, pur rimanendo una fonte consistente del lessico italiano, perderà d'importanza. Il XVIII secolo rappresenta invece il periodo con maggiori entrate nel lessico da lingue straniere e nei secoli a venire gli apporti stranieri supereranno quelli latini che andranno via via riducendosi senza tuttavia scomparire.

Secolo	Latino (%)	Italiano (%)	Altre lingue (%)
XV	45,4	42,4	12,1
XVI	38,7	43,8	17,5
XVIII	18,8	45,0	36,1
XX	3,7	70,6	25,7

Tabella 2: Percentuale di parole di formazione latina, italiana e altre lingue per secolo.

1.8. Il Lessico Elementare

Tra la fine degli anni Ottanta e gli inizi degli anni Novanta, alcuni studiosi hanno condotto una ricerca tesa a individuare quali sono le parole conosciute dai bambini; dai risultati del loro studio hanno elaborato un dizionario che è stato chiamato Lessico Elementare.

Gli autori (Lucia Marconi, Michela Ott, Elia Pesenti, Daniela Ratti, Mauro Tavella) si sono chiesti quali e quante fossero le parole conosciute dai bambini e hanno voluto creare un dizionario aggiornato e completo con una duplice finalità: da un lato costituire un punto di partenza per ricerche linguistiche, dall'altro fungere da strumento di lavoro per coloro che si occupano dei bambini e della loro educazione. L'indagine, infatti, anche se principalmente linguistica, ha risvolti anche di tipo sociale, psicologico e cognitivo.

Il Lessico Elementare dunque è, come si legge nell'*Introduzione* al volume, «un lessico di frequenza della lingua scritta, cioè un dizionario che contiene una lista di lemmi tratti da testi scritti, scelti in base alla loro frequenza d'uso» (Marconi *et al.* 1994, p. 7) e contiene le 6095 parole più frequentemente usate nei testi scritti da e per i bambini.

Nel volume sono presenti anche altre due liste di frequenza: il Lessico di Lettura, che contiene le parole che i bambini leggono di più, e il Lessico di Scrittura, che al contrario contiene le parole che i bambini scrivono di più.

La ricerca è stata svolta a livello nazionale, quindi con un campione ampio ed eterogeneo, durante il triennio 1987-1989, periodo al quale si riferiscono i materiali usati per il corpus. Il Lessico di Scrittura è stato tratto da testi scritti

dagli alunni delle cinque classi elementari, mentre il Lessico di Lettura è stato tratto dai materiali che i bambini leggono di più, cioè testi scolastici, libri di lettura e di fiabe, giornalini e fumetti.

Questo studio risulta ancora più interessante se si considera che non esistono per l'italiano altre analisi del lessico di frequenza dei bambini. Anche l'inglese, che è la lingua con il maggior numero di contributi specifici sul lessico dei bambini, ha un numero esiguo di lavori sull'argomento.

Il corpus di riferimento del Lessico Elementare è stato selezionato in ugual misura tra i testi prodotti per i bambini e quelli scritti dai bambini stessi, per un totale di 1.000.000 di occorrenze, 500.000 ciascuno.

Per le 'parole lette' il panorama di letture rivolte ai bambini è stato suddiviso in tre sottocategorie: libri di lettura e fiabe (L), giornalini e fumetti (F) e libri di testo (T). I ricercatori hanno agito in questo modo poiché «è stato valutato che questi tre settori rappresentino abbastanza fedelmente il panorama delle letture rivolte ai bambini della scuola elementare» (Marconi *et al.* 1994, p.11). Per ognuna delle tre sottocategorie sono stati scelti 100 testi e da ognuno di questi sono state estratte 1700 parole, quindi 170.000 parole per ogni sottocategoria, per un totale di circa 500.000 occorrenze.

La scelta dei libri di lettura è stata affidata a una statistica nazionale realizzata da Demoskopea, mentre per selezionare i libri di testo si è ricorsi a statistiche non pubblicate. Per quanto riguarda invece i giornalini e i fumetti sono stati scelti i più diffusi negli anni 1987-1989 e tra questi si è cercato di rappresentare una varietà di argomenti.

Per le 'parole scritte' il campione è stato costruito tenendo conto delle classi scolastiche e del sesso e sono stati presi in considerazione testi scritti da bambini di tutte le regioni d'Italia. In totale sono stati analizzati 5000 temi di lunghezza diversa suddivisi nelle cinque classi. È stata inoltre rispettata la proporzionalità rispetto alla popolazione scolastica divisa per regione. In base alle variabili prese in considerazione (classe, sesso, provenienza), sono state fatte le seguenti suddivisioni: sono stati considerati 1000 temi per ogni classe, di cui metà scritti da femmine e metà da maschi e per ogni regione metà degli elaborati è stata presa

dai centri capoluogo di provincia e metà da centri fuori dal capoluogo. I ricercatori non hanno assegnato un argomento, ma hanno delegato la scelta agli insegnanti o direttamente ai bambini, in modo da ottenere un corpus quanto più possibile diversificato al suo interno. Anche la lunghezza dei testi non è stata stabilita a priori, ogni bambino poteva scrivere un testo della lunghezza desiderata in base alla sua volontà e alle sue capacità. Va inoltre sottolineato che i testi non sono stati corretti dai docenti.

Per avere un quadro di riferimento più completo, agli insegnanti è stato chiesto di compilare una scheda anagrafica per ogni bambino, in modo da ricavare informazioni sul suo ambiente socio-culturale e sul contesto territoriale della scuola. Accanto a queste informazioni è stato chiesto anche di indicare la professione e il grado di istruzione dei genitori. Tutte queste informazioni collaterali, pur non influenzando la scelta del campione, sono servite ai ricercatori per inserire i dati ricavati dalla ricerca in un contesto più dettagliato. È, infatti, risaputo che gli stimoli ambientali sono in grado di influire fortemente sulle conoscenze lessicali della persona: un ambiente culturalmente stimolante e diversificato al suo interno favorisce l'apprendimento e l'uso di un numero maggiore di vocaboli.

Per la realizzazione del Lessico Elementare il lavoro è stato suddiviso in due fasi: in un primo momento si è dovuto reperire e catalogare i materiali presi in esame, poi, in un secondo momento è stato necessario immettere i dati, lemmatizzare in modo semiautomatico le parole estrapolate dal corpus, disambiguare gli omografi e infine effettuare l'analisi statistica. Durante la prima fase, dopo aver reperito i testi scritti dai bambini, si trattava di decidere come considerare gli errori. Gli autori hanno pertanto stabilito che (Marconi *et al.* 1994, p. 15):

Nel rispetto dell'obiettivo che ci eravamo proposti, quello cioè di produrre un lessico della lingua scritta, conosciuta e prodotta dai bambini delle elementari, [abbiamo deciso] di usare il criterio di correttezza ortografica come criterio base per inserire i singoli termini nel corpus di riferimento. Tutte le parole ortograficamente corrette sono state pertanto acquisite, indipendentemente dalla loro correttezza d'uso dal punto di vista morfologico, sintattico e semantico, mentre sono state scartate tutte quelle che non rispettavano le convenzioni ortografiche della lingua italiana.

Dopo l'inserimento dei dati si sono usati dei software per la lemmatizzazione⁴, cioè per associare a ciascuna parola il lemma o i lemmi corrispondenti e per assegnare ad ogni lemma la categoria grammaticale di appartenenza. È stato comunque necessario intervenire manualmente per disambiguare gli omografi.

Per l'elaborazione statistica dei dati sono stati identificati gli strumenti statistici e matematici adatti a verificare le ipotesi di partenza. Infatti, considerata la frammentazione storico-culturale dell'Italia, gli autori hanno scelto i parametri di classe, sesso, regione e capoluogo di provincia, per rispondere alle seguenti ipotesi (Marconi *et al.* 1994, pp. 19-20):

Il parametro classe, in quanto si presuppone che la conoscenza del lessico si incrementi con l'età e il livello di istruzione; il parametro sesso, in quanto si presuppone che l'apprendimento del lessico possa essere influenzato da interessi e contesti culturali diversi per i due sessi; il parametro regione, in quanto l'unità nazionale, da un punto di vista linguistico, è un fatto ancora recente che può far presupporre l'influenza dialettale sull'apprendimento del lessico; il parametro capoluogo – non capoluogo, in quanto l'industrializzazione e la conseguente urbanizzazione, da un punto di vista linguistico, sono fenomeni recenti.

Durante l'elaborazione statistica, i ricercatori non si sono fermati alla frequenza assoluta delle parole, nella consapevolezza che un vocabolo può apparire molte volte in un testo ma mai negli altri. Per correggere la frequenza assoluta è stato dunque calcolato l'indice di dispersione, seguendo l'esempio degli altri lessici di frequenza. L'indice di dispersione è, ricordiamolo, (Marconi *et al.* 1994, p. 23):

La misura della presenza del lemma nei diversi sottosettori dell'intero corpus ed è calcolato in modo tale che le parole con uguale frequenza assoluta totale che tendono a concentrarsi in pochi sottosettori risultino con frequenza relativa più bassa rispetto alle parole diffuse in più sottosettori.

In questo modo se un lemma è molto diffuso in molti testi diversi il suo indice di dispersione si avvicinerà a 1, nel caso contrario sarà più prossimo a 0. L'indice di dispersione è dunque fondamentale per capire quali sono le parole realmente più frequenti.

⁴ La 'lemmatizzazione' nel *Lessico Elementare* (Marconi *et al.* 1994, p.16) viene anche definita «come l'operazione, automatica o manuale, con cui si riconduce ciascuna occorrenza o parola di un testo al suo lemma fondamentale o entrata di dizionario, individuandone nel contempo una serie di informazioni (cat. grammaticale, genere, etc.) che variano secondo gli obiettivi prestabiliti».

Dallo spoglio dell'intero corpus sono risultati 19.624 lemmi diversi e da questi, sulla base delle analisi statistiche effettuate, è stato prodotto il Lessico Elementare che, come già ricordato, contiene i 6095 lemmi più frequentemente usati nei testi scritti da e per i bambini. In una lista separata sono stati anche riportati i 495 nomi propri più frequenti.

Nell'*Introduzione*, per approfondire maggiormente l'indagine, il lessico è stato suddiviso in classi di frequenza ed è risultato che i primi 500 lemmi coprono circa l'80% delle occorrenze (nel Lessico di Scrittura il dato sale all'84% del totale, evidenziando una rilevanza ancora maggiore). È risultato inoltre che le prime 1405 parole del Lessico coprono il 90% dell'intero vocabolario utilizzato. L'andamento complessivo risulta comunque analogo a quello del LIF (i primi 1000 lemmi del LIF coprono l'87% delle occorrenze). Le parole 'necessarie' ai bambini per esprimersi in forma scritta sono dunque poche: con 752 lemmi realizzano il 90% dei loro testi.

Se ci si chiede invece quali siano le categorie grammaticali più usate, la categoria più significativa risulta essere quella dei sostantivi i quali rappresentano il 54% dell'intero Lessico. Verbi e aggettivi seguono rispettivamente con il 21% e 17%. Diversamente, se si considerano le occorrenze anziché i lemmi, si nota che, come è logico aspettarsi, le parole grammaticali sono le più usate; dunque (Marconi *et al.* 1994, p.33):

Per scrivere e per leggere i sostantivi, i verbi e gli aggettivi che si devono conoscere sono molti di più rispetto alle parole grammaticali, ma queste ultime sono percentualmente molto più utilizzate (anche se sono molto minori in numero e sempre le stesse).

Confrontando i tre Lessici, quello Elementare, quello di Lettura e quello di Scrittura, si nota che quasi tutte le parole del Lessico di Scrittura sono comprese nel Lessico Elementare, mentre non si può dire lo stesso per quello di Lettura che contiene molte parole che non sono comprese nei 6095 lemmi del Lessico Elementare. Il Lessico di Lettura è, come è lecito aspettarsi, molto più ampio rispetto a quello di Scrittura e molti sono quindi i termini esclusivi di questo lessico. Non ci potrebbe, infatti, essere incremento delle competenze lessicali se i libri per bambini usassero solo le parole che già conoscono e usano. Quindi

soltanto il Lessico di Scrittura può essere visto come un sottoinsieme del Lessico Elementare.

Tuttavia ci sono lemmi che sono usati dai bambini ma che non compaiono nei testi scritti per loro (76 in totale); tra questi, alcuni sono entrati a far parte del Lessico Elementare, mentre altri, meno frequenti, non vi compaiono. I lemmi sono significativi soprattutto dal punto di vista semantico, poiché è sorprendente trovare nel testo di un bambino parole come *siringa, drogato, depuratore, ettaro, fosfato, mafia, capoluogo*, etc. «Sembra quindi» osano concludere gli autori, «che il mondo degli scritti dei bambini sia già più avanti, più al passo con i tempi, rispetto a quello delle loro letture» (Marconi *et al.* 1994, p.35).

Sicuramente non bisogna dimenticare che la ricerca qui brevemente esposta si riferisce a un periodo storico preciso (tra la fine degli anni Ottanta e l'inizio dei Novanta), ma, per quanto sia doveroso tener presente dei cambiamenti che la società, e in conseguenza anche la lingua, ha subito negli ultimi 25 anni, è lecito considerare il Lessico Elementare come una base di partenza tutt'ora attendibile per ricerche in campo linguistico.

2. Lo sviluppo lessicale nei bambini e come misurarlo

2.1. Cenni sulle teorie di acquisizione del linguaggio

La prima parola pronunciata da un bambino è un evento ricordato con gioia dai genitori di tutto il mondo, è il giorno in cui il bambino comincia il suo straordinario percorso che lo porterà ad acquisire in breve tempo il codice linguistico usato dei genitori e dal resto della comunità linguistica di cui fa parte. Scoprire quali meccanismi sottostanno e supportano il processo di acquisizione del linguaggio è un punto su cui si sono interrogati linguisti e psicologi giungendo a formulare diverse teorie. Ripercorriamole ora brevemente.

Alla fine degli anni Cinquanta viene elaborata da Skinner la teoria comportamentista secondo la quale l'apprendimento del linguaggio avviene per imitazione, rinforzo e punizione. In questo senso, l'apprendimento del linguaggio sarebbe del tutto simile all'apprendimento di qualsiasi altro comportamento. Secondo questa teoria l'ambiente riveste un ruolo fondamentale: gli enunciati emessi dal bambino vengono scoraggiati se scorretti, rinforzati se adeguati. In questo modo il bambino raccoglie informazioni riguardo alla sua lingua e modifica le sue produzioni per conformarsi sempre più a quelle degli adulti.

A due anni di distanza dalla pubblicazione dell'opera di Skinner, Chomsky ne fa una recensione ed evidenzia i limiti della teoria comportamentista. I punti centrali dell'argomentazione chomskiana sono principalmente due: viene constatato che la teoria comportamentista non è in grado di spiegare né la rapidità e l'ampiezza quantitativa dell'acquisizione del linguaggio (tramite il sistema stimolo-risposta-rinforzo, l'acquisizione del linguaggio avverrebbe in tempi pari quasi alla durata della vita umana), né l'uso creativo che il bambino è in grado di fare delle informazioni che possiede sulla sua lingua (il bambino può infatti produrre frasi che non ha mai sentito prima, fenomeno del tutto inspiegabile dalla teoria di Skinner). La posizione di Chomsky è infatti opposta a quella comportamentista: se, come abbiamo detto, questa teoria ritiene centrale il ruolo svolto dall'ambiente, la teoria innatista di Chomsky ipotizza l'esistenza di un meccanismo specie-specifico innato il *Language Acquisition Device* (LAD). Questo renderebbe

conto del fatto che tutti i bambini, indipendentemente dalla lingua che devono apprendere, dalla qualità e dalla quantità dello stimolo linguistico proveniente dagli adulti, apprendono il linguaggio con uguale rapidità e seguendo le stesse tappe. Questa concezione privilegia il ruolo della sintassi lasciando al lessico un ruolo marginale: appresa la struttura profonda, la 'grammatica', basta soltanto applicare le etichette lessicali per ottenere l'enunciato (struttura superficiale). Solo a partire dagli anni Settanta le prospettive semantiche riacquisiscono importanza, anche se dal punto di vista dell'acquisizione sintassi e semantica differiscono notevolmente: se la prima può dirsi acquisita per la maggior parte già a quattro o cinque anni, la seconda ricopre un arco di tempo molto più ampio e non si conclude neanche con il raggiungimento dell'età adulta.

In parte collegata al filone innatista è la teoria di Piaget secondo la quale quando il neonato viene al mondo egli possiede una struttura di partenza sulla quale innestare le sue conoscenze, tra cui anche quelle relative al linguaggio. Al centro della sua teoria stanno, infatti, i processi cognitivi nella loro globalità, e il linguaggio è soltanto uno di questi processi. Il raggiungimento di un certo livello nello sviluppo cognitivo è la condizione necessaria affinché possa avviarsi la funzione simbolica dalla quale dipende il linguaggio, «c'è quindi un rapporto di interdipendenza tra lo sviluppo cognitivo, lo sviluppo di una più ampia funzione simbolica [...] e i primi stadi dell'acquisizione del linguaggio» (Basile 2012, p. 89).

Tra le posizioni innatiste e quelle comportamentiste si collocano le teorie dell'interazionismo sociale che sottolineano l'importanza degli scambi comunicativi tra il bambino e gli adulti. Non rifiutano l'ipotesi dell'esistenza di un LAD ma ritengono che sia di fondamentale importanza sostenere attivamente il progressivo emergere del sistema linguistico nel bambino da parte dell'adulto. I ricercatori che si inseriscono in questo filone di pensiero, tra i quali, solo per fare un esempio, Bruner, rivolgono la loro attenzione, oltre che alle produzioni del bambino, anche alle caratteristiche del linguaggio usato dagli adulti quando devono comunicare con i bambini. In questa teoria, dunque, sia la dotazione genetica che quella ambientale ed esperienziale hanno pari importanza.

2.2. Lo sviluppo lessicale nella terza infanzia⁵

Il bambino impara a parlare rapidamente e senza sforzo, semplicemente perché è esposto all'input linguistico proveniente dagli adulti. L'ascolto e la ricezione sono infatti fasi di fondamentale importanza che precedono e preparano il bambino alla produzione linguistica: egli immagazzina le informazioni che riceve dall'ambiente circostante, le collega tra di loro e comincia ad elaborare una sua grammatica. Dunque anche prima di cominciare a parlare il bambino capisce un gran numero di vocaboli. Poco dopo il suo primo compleanno, invece, il bambino pronuncia in media otto parole isolate, costituite soprattutto da due sillabe ripetute, ognuna delle quali formata da una consonante e una vocale (es. *mama, papa, tata*, etc.). Inoltre, il bambino tende a fare un uso olofrastico delle parole, cioè con un'unica parola sottintende un enunciato più complesso (una constatazione, una domanda, una richiesta). A due anni e mezzo è invece in grado di comporre brevissime frasi formate da due parole e conosce circa 400 vocaboli: questo «è certamente il periodo in cui si registra una crescita esponenziale di vocaboli nuovi» (Ferreri 2005, p.16). C'è dunque una correlazione tra età ed estensione del vocabolario: con l'ampliarsi degli orizzonti sociali e con l'inizio del percorso scolastico, il vocabolario del bambino si espande fortemente, grazie anche alla lettura, così a cinque-sei anni conosce circa tre-quattromila parole con un incremento annuo che si aggira tra i due e i tremilacinquecento vocaboli. Durante l'adolescenza le parole note sono già decine di migliaia e alla fine delle scuole superiori mediamente sono conosciute 40.000 parole. In base poi al grado di istruzione, agli interessi personali e alle necessità legate all'attività lavorativa, un adulto può arrivare a conoscere più di ottanta o anche centomila parole. A differenza della grammatica, infatti, il lessico non smette mai di ampliarsi e modificarsi nella mente del parlante, specularmente a quanto avviene nel sistema linguistico: la natura aperta e variabile del lessico si ritrova anche nel singolo parlante. Va tuttavia specificato che i dati numerici sono soggetti a consistenti variazioni da individuo a individuo e che per l'italiano i dati disponibili sono parziali, a differenza di quanto avviene

⁵ Con 'terza infanzia' ci si riferisce a quel periodo di vita compreso tra i sei e i dieci anni di età.

invece per l'inglese. «L'estensione del vocabolario italiano nella fascia dai 6 ai 18 anni non è in Italia oggetto di indagini campionarie sistematiche né nel tempo né nello spazio» (Ferreri 2005, p.31): i dati provengono perlopiù da studi condotti in singole realtà, utili per dare un'idea approssimativa ma non adatti a fare considerazioni più ampie. Tuttavia è stato notato che le parole che i bambini italiani pronunciano per prime sono nomi e questo rinvia sia all'attività di denominazione che alla predominanza di questa categoria nel lessico. Con l'aumentare dell'età però, le altre categorie grammaticali incrementano il loro peso nella composizione del vocabolario. Va precisato tuttavia che il predominio del nome nella prima fase di acquisizione della lingua, pur caratterizzando diverse lingue, non viene considerato un universale da tutti gli studiosi. Recenti studi condotti sul cinese e il coreano hanno infatti messo in luce una prevalenza dei verbi sui nomi negli enunciati dei bambini parlanti queste lingue.

In ogni caso, passata questa prima fase di acquisizione linguistica, durante gli anni di scuola primaria il bambino vive un rapido incremento lessicale che però, se confrontato con quello sintattico, risulta essere lento. Se, come abbiamo già avuto modo di ricordare, a quattro-cinque anni l'organizzazione sintattica può dirsi completata, quella lessicale no. Per spiegare questo fenomeno sono state avanzate diverse ipotesi. Alcuni sostengono che lo sviluppo di una certa competenza lessicale sia strettamente legata al raggiungimento di un corrispettivo livello di maturità intellettuale: i sostenitori della teoria di Piaget, infatti, vedono nella competenza linguistica un riflesso dei cambiamenti cognitivi in corso tra i 5 e i 14 anni. Altri attribuiscono questa lentezza alla complessità dell'informazione codificata nel lessico, mentre un'altra spiegazione si riferisce all'astrattezza dei tratti semantici: nella forma fonica e grafica di una parola non ci sono indizi che rimandano direttamente al suo significato, la forma superficiale non fornisce indicazioni riguardo una possibile regolarità semantica sottostante. Alcuni studiosi hanno anche suggerito che l'accrescimento del lessico avviene seguendo due fasi distinte: il bambino prima assegna la parola nuova con cui entra in contatto a un'area semantica e poi, in un secondo momento, individua le differenze che

intercorrono tra questa e le altre parole che fanno parte della stessa area semantica.

Le difficoltà legate alle conoscenze lessicali aumentano se si considerano quelle parole che oltre al significato letterale hanno anche un significato traslato-psicologico, come per esempio *dolce* che si può riferire sia al gusto di qualcosa che a una persona. In questi casi i bambini considerano le due forme omonime come vocaboli distinti che si riferiscono a referenti diversi e solo tra i 10 e i 12 anni comprendono che si tratta di un'unica forma con due significati, individuando così la relazione semantica che li lega. Difatti, pur comprendendo il significato della frase "siamo circondati da persone brillanti", non riescono a individuare un collegamento tra questo uso della parola *brillante* e quello che compare ad esempio in "quel vestito ha colori brillanti". La comprensione e l'uso di vocaboli polisemici richiedono dunque elevate competenze linguistiche e hanno bisogno di più tempo per maturare. Questo fenomeno è strettamente legato al fatto che fino ai 10 anni le parole conosciute dai bambini sono perlopiù concrete, mentre il passaggio verso una concettualizzazione più astratta, necessaria per collegare il significato fisico a quello psicologico, avviene più tardi.

Negli ultimi anni di scuola primaria si osserva un nuovo modo di approcciarsi alla lingua: essa diventa oggetto di riflessione autonomo, svincolato dalla realtà extralinguistica. In questa fascia d'età quindi, con lo sviluppo della riflessione metalinguistica e della capacità di astrazione e generalizzazione, il bambino comincia a manifestare una maturità linguistica simile a quella degli adulti. In terza queste nuove strategie di riflessione sono solo abbozzate, mentre in quinta si manifestano pienamente.

2.3. Gli obiettivi stabiliti dal MIUR⁶

In ogni momento della vita di un parlante il vocabolario che egli capisce e usa può cambiare e aumentare e non si può trascurare l'importanza che il lessico ricopre nella vita di una persona, di qualsiasi età. Possedere le parole adeguate per

⁶ Ministero dell'Istruzione, dell'Università e della Ricerca.

esprimere i concetti che si vogliono veicolare è di fondamentale importanza per la vita in società, come sottolinea efficacemente Tullio De Mauro (2003, p.111):

Ricordiamoci che se vogliamo fare il mestiere più difficile, il mestiere di esseri umani e persone civili, possiamo e dobbiamo trovare, fra le parole della lingua, quelle che fanno viaggiare meglio i sensi che vogliamo esprimere.

Le conoscenze lessicali derivano in molta parte dalle esperienze personali e dal contesto socio-culturale nel quale la persona vive e certamente tutto ciò influenza profondamente anche lo sviluppo lessicale del bambino, che come gli adulti risente degli input provenienti dall'ambiente circostante. Ma non si può negare che la scuola ricopra un ruolo di grande importanza nella vita del bambino. La scuola, infatti, non solo è il luogo nel quale si scoprono le diverse aree del sapere, ma è anche il luogo nel quale si concentrano molte delle esperienze del bambino e della sua vita sociale. È per questo motivo che anche, e soprattutto, la scuola deve porsi tra i suoi obiettivi quello di incrementare le conoscenze lessicali dei bambini e dei ragazzi.

Nelle *Indicazioni nazionali per il curricolo della scuola dell'infanzia e del primo ciclo d'istruzione* divulgate dal MIUR nel settembre 2012, vengono affrontati molti temi centrali per l'istruzione e tra questi anche quello delle competenze lessicali (compresi gli obiettivi da raggiungere). Innanzitutto viene specificato che (p.9):

Nella consapevolezza della relazione che unisce cultura, scuola e persona, la finalità generale della scuola è lo sviluppo armonico e integrale della persona, all'interno dei principi della Costituzione italiana e della tradizione culturale europea, nella promozione della conoscenza e nel rispetto e nella valorizzazione delle diversità individuali, con il coinvolgimento attivo degli studenti e delle famiglie.

In secondo luogo viene sottolineato che se «lo Stato stabilisce le norme generali cui devono attenersi tutte le scuole, [fissando] gli obiettivi generali, gli obiettivi di apprendimento e i relativi traguardi per lo sviluppo delle competenze dei bambini e ragazzi per ciascuna disciplina campo di esperienza» (p. 9), sono le scuole che elaborano autonomamente il proprio curriculum.

Coerentemente con quanto stabilito dal Parlamento europeo e dal Consiglio dell'Unione Europea le competenze-chiave per l'apprendimento sono le seguenti:

1. Comunicazione nella madrelingua

2. Comunicazione nelle lingue straniere
3. Competenza matematica e competenze di base in scienze e tecnologia
4. Competenza digitale
5. Imparare a imparare
6. Competenze sociali e civiche
7. Spirito di iniziativa e imprenditorialità
8. Consapevolezza ed espressione culturale

In particolare, per quanto riguarda la madrelingua viene specificato come la competenza linguistica sia di imprescindibile importanza «per la crescita della persona e per l'esercizio pieno della cittadinanza, per l'accesso critico a tutti gli ambienti culturali e per il raggiungimento del successo scolastico in ogni settore di studio» (p. 28).

La situazione italiana, a differenza degli altri paesi europei, presenta per ragioni storiche una situazione linguistica particolare. Il ritardo nel raggiungimento dell'unità politica ha portato con sé, ovviamente, anche un ritardo dell'unità linguistica che solo negli ultimi anni sta raggiungendo il compimento. In Italia, infatti, oltre all'italiano, sono parlate molte lingue minoritarie (tutelate dalla legge n. 482/1999), lingue straniere parlate perlopiù dalle comunità immigrate e svariati dialetti; questi ultimi rappresentano tutt'oggi una realtà che non si può ignorare, anche se contano sempre meno parlanti.

Da tutto ciò si evince che ogni bambino, quando comincia la scuola primaria, possiede un patrimonio lessicale peculiare, dato dall'ambiente di provenienza e dalle sue esperienze pregresse. Ogni insegnante dovrà quindi innanzitutto consolidare la conoscenza del vocabolario di base ed espandere poi il patrimonio lessicale attingendo ai lessici delle discipline di studio. Sviluppare le competenze lessicali non è compito, infatti, del solo insegnante di italiano, ma è un obiettivo condiviso da tutti gli insegnanti, indipendentemente dalla disciplina che insegnano.

L'attenzione degli insegnanti dovrà rivolgersi all'espansione sia del lessico ricettivo (quello compreso) che di quello produttivo (lessico effettivamente usato), tenendo conto della specificità del contesto comunicativo in cui la parola si inserisce. Oltre

alla competenza linguistica non andrà trascurata, infatti, la competenza comunicativa, cioè la capacità di adeguare il proprio linguaggio al contesto comunicativo nel quale l'atto linguistico si inserisce. Chiunque può verificare in base alla propria esperienza che una stessa frase inerita in contesti diversi assume significati diversi; questo avviene perché emittente, destinatario, scopo della comunicazione, contesto sociale, etc. influiscono in modo determinante sul significato dell'enunciato.

Nelle *Indicazioni* si consiglia, inoltre, di insegnare fin dalla scuola primaria a utilizzare correttamente i dizionari, strumenti utili per ampliare il numero di parole conosciute e per specificare meglio il loro significato.

Al termine della scuola primaria, dunque, il bambino dovrebbe comprendere e usare in modo adeguato il vocabolario fondamentale e quello ad alto uso sia nella forma scritta che in quella orale. Dovrebbe inoltre: riconoscere le principali relazioni di significato tra le parole (sinonimie, antonimie e campi semantici), riconoscere l'accezione con cui una parola viene usata in un particolare contesto, comprendere e utilizzare parole specifiche legate alle discipline studiate. Tuttavia, oltre a queste indicazioni generali, il MIUR lascia molta libertà d'azione ai singoli istituti e insegnanti, senza chiarire ulteriormente strumenti e metodi per raggiungere gli obiettivi prefissati.

Della valutazione del sistema scolastico, e quindi delle competenze di alunni e studenti, si occupa invece l'INVALSI (Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione). È un ente di ricerca che ogni anno effettua verifiche sulle conoscenze degli studenti tramite prove strutturate a cui vengono sottoposti gli alunni delle classi seconda e quinta primaria, prima e terza secondaria di primo grado e della classe seconda delle scuole secondarie di secondo grado. In particolare si testano le conoscenze di italiano e matematica nei vari livelli di scolarità.

Per quanto riguarda la prova di italiano per la classe quinta della scuola primaria, il test è diviso in due parti: la prima parte è dedicata alla comprensione dei testi scritti, la seconda alla riflessione sulla lingua e alle competenze grammaticali. Anche le competenze lessicali vengono valutate nelle prove che spesso

contengono quesiti nei quali si richiede all'alunno di spiegare il significato di una parola inserita in un contesto frasale o testuale. Oltre alla conoscenza lessicale, quindi, si valuta anche la capacità di ricavare il significato di una parola servendosi del contesto. Vi sono anche quesiti in cui si richiede di trovare l'intruso tra quattro sinonimi o tra parole derivate da uno stesso lemma, valutando così sia le conoscenze morfologiche, che le abilità di stabilire relazioni di significato.

2.4. I test lessicali

Oltre all'INVALSI anche i ricercatori, con finalità ovviamente differenti, hanno sentito la necessità di ideare dei test per ricavare informazioni, ma prima di vedere quali sono i test lessicali e quali informazioni ci forniscono, bisogna chiarire che cosa si intende per padronanza linguistica. Questa espressione può essere definita come un «modello globale di conoscenza, o competenza (*competence*) negli usi di una lingua, che non tiene conto di come, dove o in quali condizioni essa è stata acquisita» (Davies 2008, p.99). A partire dagli anni '70 il significato di 'padronanza' è stato ampliato fino a comprendere globalmente l'abilità linguistica, per giungere oggi a sovrapporsi in parte alla nozione di competenza (*competence*) comunicativa. Per descrivere la padronanza linguistica sono state dunque realizzate delle scale di padronanza; queste individuano dei livelli, costruiti arbitrariamente, per ogni abilità (lettura, scrittura, ascolto e produzione orale). Per valutare a che livello si colloca un apprendente, si possono usare test di padronanza, che "misurano" appunto la lingua appresa da un discente.

Considerando in particolare la conoscenza del lessico, le diverse tipologie di test si collocano in un *continuum* che va dai test più espliciti, a punti discreti (che testano quindi la singola parola presa isolatamente), a quelli più impliciti nei quali sono richieste conoscenze più ampie per riuscire a svolgere il compito (come nei *cloze test*). Il presupposto teorico di partenza prevede che il significato di una parola sia legato al contesto frasale nel quale quella parola è inserita, quindi, estremizzando, le parole prese isolatamente sono prive di significato.

Nel saggio di Alan Davies contenuto nel volume della collana GISCEL⁷ *Lessico e apprendimenti. Il ruolo del lessico nella linguistica educativa* vengono proposti diversi test lessicali. Ripercorriamoli ora brevemente dal più esplicito al meno esplicito.

Nella prima tipologia viene chiesto di segnare quale delle parole sotto riportate ha il significato più vicino alla parola target ("quale delle seguenti quattro scelte ha il significato più vicino alla parola 'x'?"). Nella seconda tipologia, dopo aver presentato la parola target all'interno di una frase, viene chiesto di scegliere quale delle spiegazioni sotto proposte corrisponde meglio al significato della parola target. Nella terza vengono presentate sei parole e tre definizioni e viene chiesto di segnare la parola corrispondente a ciascuna definizione. Nella quarta vengono fornite delle frasi, in ognuna delle quali una parola è incompleta; viene chiesto di completarla in modo da rendere le frasi di senso compiuto. Nella quinta tipologia, simile alla quarta, le frasi proposte compongono un testo e viene richiesto di completare le parole in modo da dimostrare di aver compreso il testo. Nell'ultima tipologia proposta da Davies viene presentato un testo in cui sono state aggiunte anche delle parole estranee al testo che andranno quindi cerciate.

Tuttavia, oltre a queste modalità, sono state sperimentate altre dai diversi gruppi GISCEL (quindi soprattutto in riferimento agli studenti). Chiara Colli Tibaldi e colleghe hanno chiesto agli alunni delle scuole primarie di dichiarare se conoscevano o meno alcune parole (soprattutto aggettivi) appartenenti alla sfera delle sensazioni tramite risposta chiusa; poi, per ogni parola dichiarata come conosciuta, hanno chiesto di scrivere una definizione e un esempio per verificare l'attendibilità della risposta fornita con la crocetta. Un lavoro simile è stato fatto anche dai ricercatori del GISCEL Calabria e del GISCEL Sicilia. I primi hanno realizzato una scheda lessicale anonima con quattro possibilità di risposta per ogni parola ("la conosco", "non la conosco", "l'ho sentita ma non la conosco", "non l'ho mai sentita") seguita da una seconda parte in cui si richiedeva di scrivere una frase per ogni parola conosciuta in modo da verificare sia l'attendibilità della risposta che l'accezione usata. I secondi hanno realizzato una ricerca scandita

⁷ Gruppo di Intervento e di Studio nel Campo dell'Educazione Linguistica.

dalle stesse fasi di quella del GISCEL Calabria, con l'unica differenza che per la fase di autovalutazione hanno utilizzato una griglia di analisi diversa, ispirata a quella proposta da Gensini e Vedovelli (1983). Per ogni parola le opzioni, infatti, erano: "parola mai sentita o letta", "sentita o letta ma di cui non si conosce il significato", "conosciuta ma non usata", "conosciuta e usata". Infine Salvatore Masi e Maria Maggio hanno proposto l'uso di vignette per eseguire diversi compiti, come riempire uno spazio vuoto di una frase servendosi dell'immagine, trovare una parola osservando una o più vignette, oppure interpretare una frase, associando ad essa la vignetta appropriata.

Come si è visto, ci sono molti modi per valutare le competenze lessicali, tuttavia la scelta non è così semplice come può sembrare. Se da un lato, come abbiamo detto, le parole devono essere inserite in un contesto per assumere significato, dall'altro più ampio è il contesto fornito e meno possiamo essere sicuri di star testando il lessico e non capacità inferenziali e di comprensione. È una sorta di paradosso che ci fa riflettere su come sia necessaria un'attenta valutazione prima di stilare un qualsiasi test che voglia indagare le competenze lessicali dei parlanti.

I processi inferenziali, infatti, si attivano spontaneamente sulla base di informazioni percettive (che vengono collegate mentalmente a situazioni sociali culturalmente determinate) e di conoscenze logiche pregresse, e consentono di ricavare nuovi significati. Il raffronto con gli schemi mentali permette, infatti, di confrontare le nuove informazioni con le conoscenze già possedute e di fare quindi una classificazione. I processi inferenziali rientrano nella dimensione pragmatica della lingua e si attivano soprattutto in presenza di situazioni sconosciute o ambigue, in modo più o meno consapevole: «la percezione di una lacuna o di una carenza di informazione spinge l'individuo verso un processo attivo di ricerca di significato» (Colli Tibaldi *et al.* 2014, p.198). Per questi motivi l'inferenza può alterare in modo significativo l'esito dei testi lessicali.

2.5. Le capacità definitorie dei bambini

Come si è appena visto, in molti test lessicali viene chiesto agli alunni di dare una definizione delle parole testate. Tutti noi abbiamo dovuto qualche volta spiegare a qualcuno il significato di una parola che conoscevamo e usavamo con disinvoltura, e ci sarà capitato di accorgerci che era più difficile del previsto. Questo perché definire significa, etimologicamente, limitare, tracciare dei limiti, che al lato pratico significa trovare le caratteristiche che distinguono una parola da un'altra. È allora chiaro come dare una definizione richieda chiarezza concettuale, capacità di discernimento, di astrazione e di generalizzazione, competenze tutt'altro che elementari e di facile acquisizione.

Dunque i bambini, non avendo ancora affinato queste capacità, utilizzano strategie definitorie in parte diverse da quelle degli adulti. Le loro definizioni sono state oggetto di studio per riuscire a riprodurre le strategie che loro usano spontaneamente nei dizionari pensati e realizzati appositamente per i bambini, di modo che, "parlando la stessa lingua", fossero più facilmente accessibili. Con l'aumentare dell'età si registra un passaggio da definizioni più concrete e legate alla situazione comunicativa in cui la definizione viene formulata, a definizioni più oggettive e razionali, in cui la situazione contingente influisce sempre meno lasciando spazio a generalizzazioni e astrazioni. Non è possibile tracciare una linea netta di demarcazione tra una modalità e l'altra, poiché ci sono ampie differenze individuali, ma si è notato che già durante gli ultimi anni di scuola primaria i bambini sanno applicare strategie definitorie più complesse, in cui la parola viene definita in relazione al resto del lessico e non facendo riferimento alla realtà extralinguistica. Questo passaggio si comincia a intravedere già tra la terza e la quarta primaria e diventa via via più evidente nel corso della quinta.

Una caratteristica molto ricorrente nelle definizioni date dai bambini è la presenza di elementi che rimandano alla loro esperienza personale. È proprio dall'esperienza personale, infatti, che spesso nascono associazioni che legano la parola a ricordi, sensazioni, situazioni tipiche che poi il bambino riversa nella definizione.

Molto frequente è inoltre il ricorso alla concretezza, anche per definire parole astratte: se si tratta di nomi i bambini tendono a usare il verbo con la stessa radice per spiegarne il significato. I verbi, infatti, riferendosi ad azioni vengono generalmente considerati più concreti dei nomi che descrivono processi o azioni. Legato in qualche modo alla concretezza è anche l'approccio descrittivo che traspare in molte definizioni: se la parola lo permette, i bambini tendono a descrivere l'oggetto o la situazione (soprattutto nel caso dei verbi) ricorrendo in particolare ai dati percettivi. Nella definizione citeranno quindi dimensioni, colori, materiali e quant'altro, in modo da restituire un "ritratto" o una "scenetta" del referente extralinguistico designato dalla parola. Talvolta del referente preferiscono invece evidenziare la funzione, usando espressioni come "serve a", "è una cosa che fa questo".

Con l'aumentare dell'età, e quindi anche delle conoscenze scolastiche, aumentano nelle definizioni i riferimenti più espliciti alle nozioni apprese. Le definizioni di questo tipo sono più strutturate e caratterizzate da un maggior grado di formalità. Al contrario, quando il bambino ritiene che sia troppo difficile dare una definizione, o quando pensa di non esser stato abbastanza chiaro, preferisce affidarsi a un esempio in grado di illustrare il significato della parola. In questo modo, più che spiegare esplicitamente il significato, lascia che questo emerga indirettamente dal contesto. Anche negli esempi si ritrovano molto spesso esperienze personali o situazioni ritenute tipiche.

Un'altra caratteristica ricorrente è l'uso di generalizzazioni, ovvero la parola da definire viene spiegata collocandola all'interno di una categoria più ampia, del tipo "il topo è un animale", o servendosi di tautologie (definizione di *coda*: "la coda di un gatto è una coda fatta di pelo").

Oltre a queste caratteristiche generali, vi sono altre peculiarità interessanti che riguardano più nello specifico la realizzazione formale delle definizioni. In molti casi, infatti, nelle definizioni troviamo vocaboli come *cosa*, *roba*, *persona*, etc. perché, essendo poveri di contenuto semantico, possono esser facilmente usati in molti contesti diversi. Per lo stesso motivo sono frequentemente usate anche espressioni come *quello che*, *uno che*, etc. In molti casi, inoltre, i bambini

ricorrono alla seconda persona singolare per esprimere un soggetto generico, sfruttando di fatto una delle possibilità messe a disposizione dalla lingua italiana per rendere la frase impersonale.

Oltre a quanto detto fin qui, nelle definizioni più mature, quelle in cui il bambino prende le distanze dalle sue esperienze personali per privilegiare gli aspetti più generali, si possono intravedere strategie definitorie più complesse e più vicine a quelle razionali di discendenza aristotelica usate, per esempio, nei dizionari. In particolare, tra la parola da definire e la definizione possono intercorrere i seguenti rapporti: di sinonimia o parasinonimia ("panciotto è un gilè"), di antonimia ("odio vuol dire non volersi bene"), di iponimia ("lettore è un signore che legge libri").⁸ A quest'ultimo tipo appartengono anche i casi in cui, più che un iperonimo, viene usato un nome generico come *cosa*, di cui abbiamo parlato prima.

2.6. Ricerche sulle competenze lessicali dei bambini

Interessi per la ricerca sul campo spesso nascono da esigenze operative e di programmazione didattica. Chi si occupa di educazione linguistica è infatti interessato a comprendere i livelli di partenza e gli obiettivi raggiungibili dagli alunni in ogni fascia d'età. Dalla commistione di esigenze didattiche e interessi teorici viene elaborato agli inizi degli anni Ottanta uno strumento d'indagine chiamato dai suoi ideatori *glotto-kit*, che mira a fornire utili spunti per vari aspetti della didattica della lingua italiana. Gensini e Vedovelli, ideatori del programma e curatori del volume che lo presenta, suggeriscono di cominciare ogni intervento di educazione linguistica dall'individuazione dei livelli linguistici di partenza e dalla raccolta di dati utili a definire il retroterra sociale e culturale di provenienza degli allievi, come professione e scolarità dei genitori e loro luogo di provenienza, abitudini di lettura, modo di impiegare il tempo libero, etc. Per verificare la ricchezza del VdB, punto di riferimento sia per le scuole primarie (in uscita da queste si stima che il vocabolario fondamentale, contenuto nel VdB, sia

⁸ Esempi tratti da Lo Duca (1986) e Biorci *et al.* (2006).

completamente posseduto sia in comprensione che in produzione) che per le secondarie di primo grado (a conclusione di queste, invece, l'intero VdB dovrebbe essere posseduto), gli autori suggeriscono di usare una griglia a quattro entrate: parole mai sentite, parole sentite ma dal significato sconosciuto, parole conosciute ma non usate, parole conosciute e usate. A questa consigliano di affiancare sempre una parte in cui si chiede di riformulare in modo autonomo le parole prese in esame. Le informazioni provenienti da una e dall'altra parte permettono di gettare luce sulle «associazioni psicologiche che la parola-stimolo [...] suscita» (Gensini e Vedovelli 1983, p.89), anche in caso di errori.

Sulla base di queste considerazioni, a breve distanza dalla pubblicazione del VdB, è stata effettuata una verifica sistematica, durata per l'intero arco dell'anno scolastico, di tutte le 6690 parole del VdB. Le rilevazioni sono state effettuate in scuole di diverso grado sparse su tutto il territorio nazionale. I risultati riportati in Vedovelli (1983), si riferiscono però ai soli dati raccolti con la collaborazione degli alunni delle classi terze delle diverse scuole secondarie di primo grado (all'epoca scuole medie). Ad ogni allievo veniva consegnata una scheda di autovalutazione con la griglia a quattro entrate sopra citata e in seguito, per accertarsi della validità delle risposte, veniva chiesto di costruire delle frasi con alcune delle parole conosciute, estratte a campione. Da questa verifica è risultato che i vocaboli appartenenti al vocabolario fondamentale erano conosciuti da più del 90% degli intervistati. È stata inoltre evidenziata una carenza conoscitiva dei lemmi riguardanti la vita civile e politica, e scarsamente conosciuti sono risultati anche i nomi di piante e fiori, le parole della vita in campagna e quelle legate al mondo dell'artigianato, i vocaboli delle esperienze percettive e sensoriali.

Sulla spinta delle riflessioni portate avanti dagli ideatori del glotto-kit ci sono state sia applicazioni operative che ricerche condotte in scuole di diverso grado. Per quanto riguarda in particolare la conoscenza del VdB, in una scuola media in provincia di Latina alcuni insegnanti hanno ritenuto utile condurre un'indagine per poter affrontare con maggior consapevolezza la programmazione didattica e quindi per offrire ai loro allievi un bagaglio lessicale fondamentale per affrontare la realtà in cui vivono (Starace 1983, pp. 343-356). Agli studenti sono state

sottoposte due schede: nella prima dovevano indicare sesso, età, classe, località di origine dei genitori e abitudini linguistiche dei genitori, nella seconda venivano proposte le parole del VdB e per ognuna gli alunni dovevano scegliere una delle seguenti opzioni: "parole mai sentite o lette", "parole sentite o lette, ma che non sapete bene che cosa vogliono dire", "parole di cui non siete ben sicuri: provate a spiegarle". In un secondo momento le parole sono state divise, oltre che per categoria grammaticale, anche per area semantica. Grazie a questo accorgimento è emerso che nell'ambito 'uomo' (comprendente vocaboli come *diffidenza*, *ipocrisia*, *ascesso* o *palpebra*) si trova il numero maggiore di parole sconosciute. Seguono per scarsa conoscenza le parole della vita pubblica (come *elettorale* e *referendum*) e della natura (come *lido* e *oasi*).

In questo filone di ricerca si inseriscono anche gli studi sulla didattica delle lingue condotti dai gruppi GISCEL di tutta Italia. In *Linguaggio, mente parole. Dall'infanzia all'adolescenza*, Salvatore De Masi e Maria Maggio hanno esposto la ricerca che hanno condotto nelle classi quarte e quinte delle scuole primarie di Lecce. Hanno elaborato un test composto da 90 parole: la maggior parte erano parole contenute scelte tra le parole del Lessico Elementare seguendo sia il criterio grammaticale che quello statistico (i 6095 lemmi del Lessico Elementare sono stati divisi in tre fasce di frequenza), gli item restanti erano parole funzione. Le parole piene che costituivano la prima parte del test erano dunque 72, equamente divise in verbi, aggettivi, sostantivi e avverbi. Hanno scelto di presentare gli item facendo uso di vignette che dovevano stimolare la produzione della parola. Per le parole funzione hanno invece presentato delle frasi che i bambini dovevano valutare come "giuste" o "sbagliate". Alla prova hanno partecipato 40 bambini estratti casualmente tra gli alunni delle classi quarta e quinta delle scuole che hanno aderito alla ricerca. Ad ogni bambino è stato anche chiesto di compilare un questionario riguardante il livello di scolarità dei genitori, la disponibilità di libri e altri supporti in casa, il rendimento scolastico. Durante l'elaborazione dei dati, queste variabili sono state messe in relazione ai punteggi ottenuti nel test ed è risultato che la variabile "scolarità del padre" influenza le competenze lessicali, così come accade per la variabile "rendimento scolastico". I

punteggi ottenuti, tuttavia, si distribuiscono approssimativamente secondo la curva normale, mostrando che la competenza lessicale è «particolarmente elevata per le parole di alta frequenza, mentre decresce per la fascia della media frequenza, per diminuire ulteriormente per quella bassa» (De Masi e Maggio 2006, p.180). Nella loro ricerca non è stato registrato un incremento significativo della competenza lessicale nel passaggio tra la classe quarta e la quinta, risultato che gli autori hanno attribuito a una mancanza di strategie didattiche esplicitamente riferite al lessico. Non hanno inoltre registrato differenze tra i sessi. Del vocabolario comune (CO) si sono occupati invece i ricercatori del GISCEL Calabria, poiché hanno notato che su questa fascia di lessico non erano ancora stati fatti studi. La ricerca ha interessato 100 parole del CO selezionate con un generatore equiprobabile di occorrenze (in grado di garantire, cioè, la stessa probabilità di estrazione per ognuno dei lemmi del CO) e ha coinvolto 848 studenti frequentanti le scuole primarie, secondarie di primo grado e secondarie di secondo grado delle cinque province calabresi. Ad ogni studente è stato chiesto di compilare una scheda lessicale anonima scegliendo tra quattro opzioni: "la conosco", "non la conosco", "l'ho sentita ma non la conosco", "non l'ho mai sentita". Conclusa questa fase, per ogni parola dichiarata conosciuta, veniva chiesto di scrivere una frase. Chi voleva poteva inoltre compilare una scheda sociolinguistica sulla professione e il titolo di studio dei genitori, sulle abitudini di lettura e sulla conoscenza delle lingue. I risultati hanno confermato la natura incrementale del lessico lungo tutto il percorso di istruzione. Una penetrazione del CO è presente già nelle scuole primarie, ma è con le scuole secondarie di primo grado che si ha un consistente salto quantitativo: il numero delle parole conosciute da almeno il 75 % degli alunni passa da 12 (primaria) a 33. Tra le scuole secondarie di primo grado e il biennio successivo, invece, l'incremento è modesto. Alla conclusione delle scuole secondarie di secondo grado le parole conosciute da almeno il 75% degli studenti sono 53, quindi poco più del 50% delle parole testate. Gli autori hanno inoltre rilevato che le parole meno conosciute sono le stesse in tutti i livelli scolastici.

Il CO, questa volta assieme al VdB, è stato oggetto di ricerca anche da parte del GISCEL Sicilia. Anche in questo caso la ricerca è stata suddivisa in due fasi: per la prima si sono serviti di un test di autovalutazione in cui per ognuna delle 100 parole da testare veniva chiesto di indicare se erano parole mai sentite o lette, sentite o lette ma di significato sconosciuto, conosciute ma non usate, conosciute e usate; nella seconda parte hanno verificato la corrispondenza tra autovalutazione e conoscenza effettiva attraverso la produzione libera di frasi contenenti le parole dichiarate come conosciute. Le parole sono state scelte casualmente all'interno di un *corpus* formato da testi curricolari di diverse discipline. Il campione era formato da studenti delle classi quinte delle scuole primarie, delle classi terze delle secondarie di primo grado e delle classi seconde delle scuole secondarie di secondo grado. La ricerca ha evidenziato una generale conoscenza sia delle parole appartenenti al VdB che di quelle del CO e uno scarto minimo tra conoscenza dichiarata e uso corretto della parola. È stata inoltre riscontrata una differenza delle accezioni usate dai ragazzi e dalle ragazze nella produzione libera di frasi. In generale hanno dunque confermato i dati di Gensini e Vedovelli (1983) per quanto riguarda il VdB e quelli del GISCEL Calabria (2008) per il CO.

Anche la ricerca svolta da Chiara Colli Tibaldi, Laura Deluigi e Michela Fraire, ha preso in considerazione sia lemmi appartenenti al CO (37) che al VdB (13). La loro indagine si è concentrata sul lessico sensoriale e in particolare le parole da testare (50) sono state scelte equamente tra i cinque sensi. Hanno partecipato alla ricerca 76 alunni delle classi terze e quinte delle scuole primarie di un paese del Piemonte. Ad ogni partecipante è stato chiesto di segnare con una crocetta su "sì" o su "no" per dichiarare se conoscevano o meno la parola presentata; per ogni parola conosciuta gli intervistati dovevano inoltre fornire sia una definizione che un esempio. Le produzioni scritte dai bambini sono state considerate valide quando dimostravano che il significato del vocabolo era stato colto. In un secondo momento agli alunni è stato sottoposto anche un *cloze* test che aveva lo scopo di indagare in che modo e in che misura i processi inferenziali possono influenzare e incrementare la competenza lessicale. Dalla ricerca è emerso come i bambini, alla

richiesta di fornire esempi e definizioni, si affidino ai linguaggi disciplinari o a strategie di ricostruzione del significato che tengono conto della forma della parola. In generale sono stati formulati più enunciati corretti per le parole che si riferivano al tatto, alla vista e all'udito. Da un punto di vista quantitativo invece è emersa una correlazione tra l'età e la competenza: i bambini di terza hanno dichiarato di conoscere tra le 2 e le 21 parole presentate, quelli di quinta tra le 13 e la totalità. È stata riscontrata una correlazione anche tra l'età e le capacità di definire e di produrre esempi e tra l'età e l'efficacia dei processi inferenziali. Nel prossimo capitolo verrà infine presentata la ricerca condotta presso le scuole primarie dell'Istituto Comprensivo "Dante Alighieri" di San Canzian d'Isonzo.

3. La ricerca sperimentale

Dallo studio della letteratura in merito è emerso che entro la conclusione del primo ciclo d'istruzione il ragazzo è in grado di padroneggiare l'intero vocabolario di base, oltre che numerose altre parole appartenenti al vocabolario comune e ai lessici specialistici.

Partendo da questo presupposto è nata la curiosità di indagare a che punto si trova questo processo durante gli ultimi anni di scuola primaria. Tra la classe terza e la classe quinta si assiste, infatti, a un interessante incremento delle competenze lessicali dei bambini, legato anche all'apporto dato dalle discipline umanistiche e scientifiche, che cominciano ad esser studiate proprio in quest'arco di tempo.

Ai fini della ricerca era tuttavia necessario concentrare l'attenzione su un ambito più specifico. Per farlo sono state compiute alcune scelte preliminari.

3.1. Le scelte preliminari

Il VdB, come si è visto, contiene parole appartenenti a tutte le classi grammaticali. Innanzitutto sono state escluse le parole a classe chiusa e tra quelle a classe aperta si è preferito circoscrivere la ricerca ai soli sostantivi che sono, non solo maggiori in numero, ma anche acquisiti prima in italiano.

È subito apparso chiaro che era necessario scegliere anche un'area semantica specifica da indagare; perciò, dopo aver selezionato tutti i sostantivi del VdB, sono stati trascritti su un foglio excel tutti i sostantivi che in qualche modo avevano a che fare con la vita politica e la partecipazione attiva alla cittadinanza. La scelta di quest'area semantica è stata guidata anche dalla particolare importanza che viene assegnata al sapere civico nelle *Indicazioni* del MIUR (cfr. § 2.3.).

All'interno di questa prima selezione sono state evidenziate le parole che comparivano anche nel Lessico Elementare. In questo modo è stato possibile calibrare più accuratamente la scelta delle parole: parole appartenenti non solo al VdB (che appunto si presume appreso completamente solo alla fine del terzo

anno di scuola secondaria di primo grado), ma anche scritte o lette dai bambini della scuola primaria.

La selezione delle parole contenute sia nel VdB che nel Lessico Elementare è stata effettuata considerando il loro rango di appartenenza nel Lessico Elementare. Poiché quest'ultimo contiene 6095 parole, esso è stato suddiviso in tre fasce di frequenza: la prima con le parole che occupano un rango compreso tra 1 e 2000 (quindi le parole più frequentemente usate nei testi scritti da e per i bambini), la seconda con le parole che occupano un rango compreso tra 2001 e 4000 e la terza con le parole con rango compreso tra 4001 e 6095.

Fatte queste considerazioni, si è potuto procedere con la selezione delle 30 parole da testare: 15 tra le parole contenute sia nel VdB che nel Lessico Elementare (5 per ogni fascia di frequenza) e 15 tra le parole appartenenti al solo VdB. Di queste 30 parole, 14 rientrano nel vocabolario fondamentale (FO), cioè *cittadino, diritto, legge, sindaco, stato, bandiera, capitale, decisione, ministro, presidente, esilio, indipendenza, comunista, fascista*, 9 nel vocabolario d'alto uso (AU) cioè *costituzione, sciopero, clandestino, connazionale, democrazia, elezione, inno, parlamento, propaganda* e 7 nel vocabolario di alta disponibilità (AD) ovvero *ideale, candidata, corruzione, emigrazione, mafia, senatrice, votazione*.

Nella scelta delle parole non si è tenuto particolarmente conto delle caratteristiche intrinseche delle parole, poiché si è preferito privilegiare, il criterio della frequenza e dell'appartenenza ai diversi vocabolari. Per questo motivo sono presenti sia parole concrete (ad es. *bandiera*) che astratte (ad es. *diritto*), sia parole animate (ad es. *cittadino*) che inanimate (ad es. *inno*).

In appendice 1 è riportata l'intera lista di parole usate per il test.

3.2. I questionari

Come si è visto in precedenza, esistono diversi modi per testare la competenza lessicale; nel nostro caso per la raccolta dei dati si è scelto di utilizzare la forma del questionario anonimo. Nel questionario ogni parola target è stata inserita in un contesto frasale limitato, poiché, se da un lato una parola presa isolatamente è

“priva di significato”, dall’altro la presenza di un contesto ampio e articolato permette quei processi inferenziali che non consentono di misurare in modo attendibile la competenza lessicale.

Dunque, dopo aver presentato la frase contenente la parola target, è stato chiesto al bambino di segnare con una crocetta se conosceva la parola, se non la conosceva o se non la conosceva ma l’aveva già sentita. Per accertare che l’indicazione data attraverso la risposta chiusa fosse attendibile, è stato chiesto di spiegare brevemente con parole proprie il significato delle parole dichiarate come conosciute.

Nel questionario veniva inoltre chiesto di indicare: la scuola e la classe di appartenenza, il sesso e l’età. La richiesta di indicare l’età serviva soltanto a confermare l’appartenenza alla classe e infatti nell’elaborazione dei dati si è tenuto conto solo della classe e del sesso. I bambini intervistati avevano dunque un’età compresa tra gli otto e gli undici anni.

3.3. La raccolta dei dati

Durante il mese di maggio 2015 sono stati somministrati i questionari a 13 classi delle scuole primarie dell’Istituto Comprensivo “Dante Alighieri” di San Canzian d’Isonzo: due classi terze, due quarte e due quinte della scuola primaria “Giosuè Carducci” di Pieris; una terza, una quarta e una quinta della scuola primaria “Giovanni Pascoli” di San Canzian d’Isonzo; due terze, una quarta e una quinta della scuola primaria “Dante Alighieri” di Turriaco. Tutte le scuole si trovano in provincia di Gorizia, a pochi chilometri di distanza l’una dall’altra. I dati raccolti dunque provengono da una realtà grosso modo omogenea e localizzata in una zona limitata.

Ogni classe aveva a disposizione un’ora per la compilazione del questionario, tuttavia, la maggior parte degli alunni ha impiegato meno del tempo a disposizione per completare il compito. Io stessa mi sono occupata della distribuzione dei questionari e della spiegazione di tutte le parti da completare, sottolineando il fatto che la prova aveva l’unico scopo di fotografare una

situazione e non giudicare chi era "più bravo" o "meno bravo". Durante la compilazione sono inoltre rimasta a disposizione degli alunni per chiarimenti circa il modo di completare le parti del questionario. Il questionario è stato distribuito a tutti i bambini, anche quelli con disturbo specifico dell'apprendimento (DSA) o bisogni educativi speciali (BES). In totale nel campione sono presenti due bambini con BES e otto con DSA. Tra i bambini del campione sono inoltre presenti alcuni di origine straniera, numericamente però poco rilevanti (5 in totale) e provenienti da famiglie che vivono in Italia da molti anni.

In totale sono stati raccolti 248 questionari così suddivisi: 95 per la classe terza, 80 per la classe quarta e 73 per la classe quinta. In particolare, i bambini che hanno partecipato alla ricerca erano 131 maschi e 117 femmine.

Alcuni insegnanti mi hanno fatto notare che l'attenzione prestata per compilare i primi quesiti poteva via via diminuire nel corso della prova; tuttavia, poiché questa è durata meno di un'ora, è possibile ipotizzare che l' "effetto stanchezza", se così si può chiamare, nel nostro caso non ha influenzato particolarmente i risultati. Comunque, tenendo conto di questo suggerimento proveniente dagli insegnanti (che, è bene sottolinearlo, conoscono meglio di chiunque altro, le capacità e le esigenze dei bambini ad ogni livello di scolarità), il numero di parole target, inizialmente fissato su quota 50, è stato diminuito a 30, e le parole che si ipotizzavano come più conosciute, e che quindi richiedevano anche una parte di produzione scritta, sono state spostate soprattutto nella prima parte del questionario.

3.4. Osservazioni durante la somministrazione

Durante i miei interventi a scuola ho potuto notare come gli atteggiamenti dei bambini nei confronti della prova siano stati molteplici. La maggior parte ha risposto ai quesiti tranquillamente e serenamente, mentre certi hanno vissuto il completamento del questionario come se si trattasse di una sorta di "verifica" (nonostante sia stato spiegato loro che si trattava soltanto di un'indagine); alcuni bambini, infatti, consegnandomi il loro questionario, hanno specificato: «non è

andata molto bene, certe parole non le conoscevo». Qualcuno, invece, si è lamentato del fatto che le parole erano "difficili".

Per quanto riguarda gli insegnanti, invece, sono stati molto disponibili e interessati alla ricerca. Alcuni hanno invitato i bambini a ragionare sulla morfologia delle parole derivate e, per quanto possibile, anche sull'etimologia. Mi hanno inoltre informata che sono stati fatti lavori di approfondimento con i bambini di quarta sulla costituzione e con quelli di terza sui diritti.

Alcuni insegnanti hanno inoltre approfittato, dopo il mio intervento, per spiegare alcune delle parole sconosciute, sulla scia di uno stimolo lanciato dai bambini stessi con le loro domande e osservazioni: c'è stato chi ha notato che la parola *emigrazione*, è simile ma non è uguale a *migrazione*; qualcun altro invece si è accorto che nonostante la parola *ideale* per lui significasse "perfetto", nella frase proposta assumeva un significato diverso, e così via.

3.5. Limiti della ricerca e possibili sviluppi

La ricerca condotta e qui esposta ha, ovviamente, dei limiti che per onestà vanno evidenziati. Innanzitutto la raccolta dei dati è stata condotta in un'area geograficamente circoscritta e socialmente abbastanza omogenea; è quindi lecito supporre che se condotta in un'altra zona, la ricerca potrebbe dare risultati parzialmente diversi. Inoltre, il campione di 248 bambini, per quanto abbastanza numeroso, non è sufficiente per consentire una generalizzazione dei risultati ottenuti. Per questi motivi, ciò che emerge dal presente lavoro è da considerarsi una tendenza.

Si aprono tuttavia possibili proseguimenti che spero verranno sviluppati in futuro. È innanzitutto possibile riproporre la ricerca in altre realtà geografiche per verificare se ci siano o meno differenze e quanto queste siano significative. È inoltre possibile estendere il campione anche ai ragazzi delle scuole secondarie di primo grado, per verificare se effettivamente a conclusione di queste il VdB è conosciuto da ogni parlante.

In conclusione dunque, questo non è che un tassello, un piccolo contributo, di una ricerca che potrebbe essere molto più ampia.

3.6. Analisi quantitativa dei dati

Dai dati raccolti tramite i questionari è emerso un notevole incremento del numero di parole conosciute tra la classe terza e la quinta, a conferma del fatto che all'aumentare dell'età aumentano anche le conoscenze lessicali. In particolare, tra la terza e la quarta si ha lo stacco maggiore, come si vede nel grafico qui sotto.

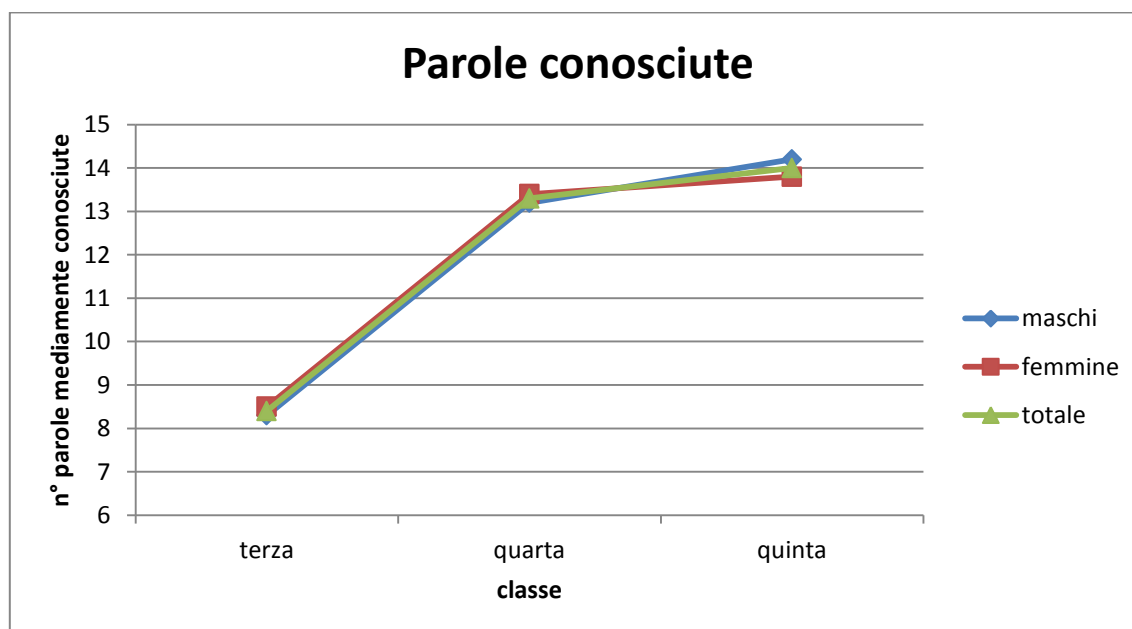


Grafico 1: Parole mediamente conosciute negli ultimi tre anni di scuola primaria.

Se, infatti, un bambino di terza primaria conosce mediamente 8,4 parole tra quelle presentate (pari al 28%), uno di quarta ne conosce 13,3 (44%) e uno di quinta 14 (47%). Il grafico, con le tre curve quasi perfettamente sovrapposte, mostra, inoltre, che l'andamento dei maschi e quello delle femmine è analogo: non ci sono quindi differenze legate al sesso.

Una sempre maggiore conoscenza delle parole testate è confermata anche dalla diminuzione del numero di parole dichiarate come non conosciute: parole non conosciute dai bambini di terza, sono dichiarate invece come conosciute o come

non conosciute ma già sentite, dai bambini più grandi. Mentre in terza le parole non conosciute sono il 30%, in quarta sono solo il 16% e in quinta il dato scende fino al 12%. La curva appare dunque ribaltata rispetto a quella delle parole conosciute.

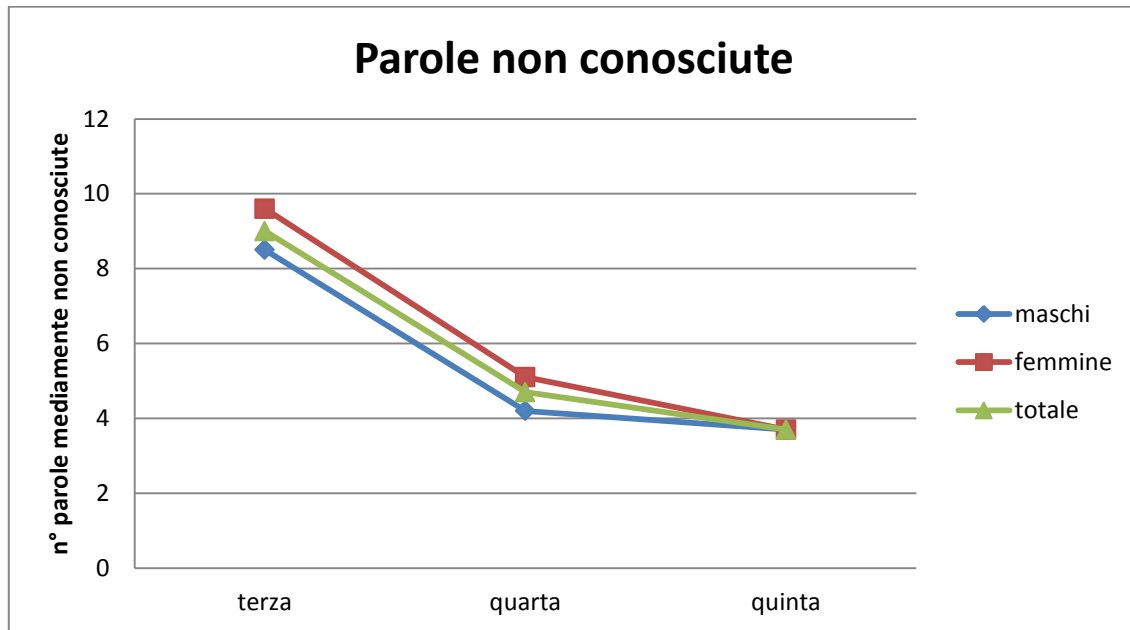


Grafico 2: Parole mediamente non conosciute negli ultimi tre anni di scuola primaria.

Anche in questo caso le tre curve tendono a sovrapporsi non evidenziando alcuna differenza significativa legata al sesso.

Prendendo in considerazione i valori medi registrati in tutte e tre le tipologie di risposta ("la conosco", "non la conosco ma l'ho già sentita" e "non la conosco"), la situazione è quella delineata dai grafici seguenti. Entrambi i grafici rappresentano il numero medio di parole dichiarate come conosciute, non conosciute e non conosciute ma già sentite, ma il primo raggruppa le risposte per opzione e permette di evidenziare come il numero di parole conosciute cresca e quello di parole non conosciute cali, mentre il secondo fotografa le tre opzioni di risposta divise per classe, mostrando quali sono state le risposte mediamente date per ogni livello.

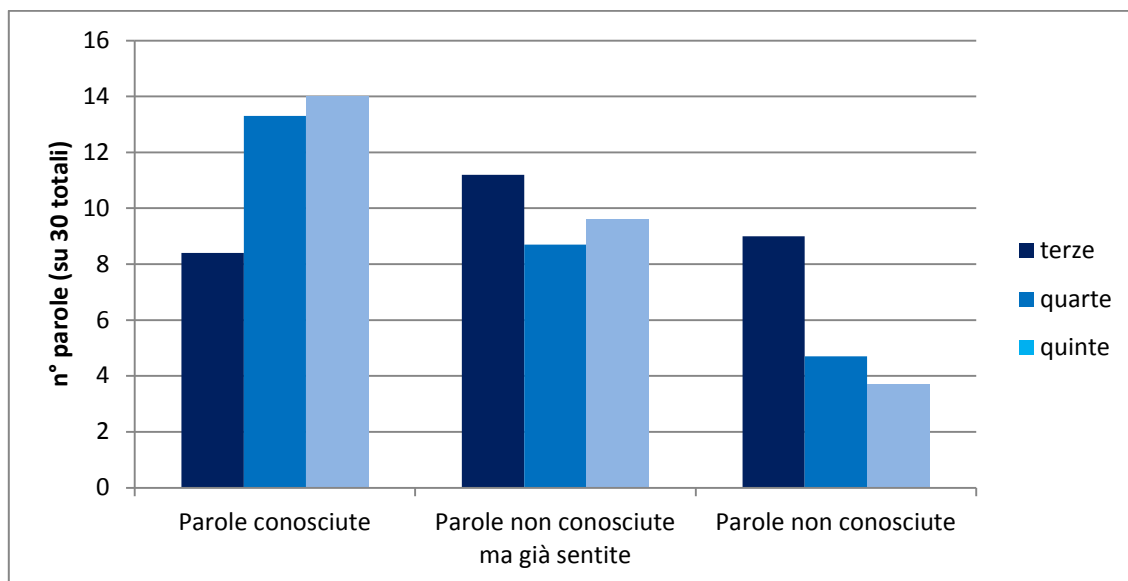


Grafico 3: Valori medi per ogni tipologia di risposta negli ultimi tre anni di scuola primaria.

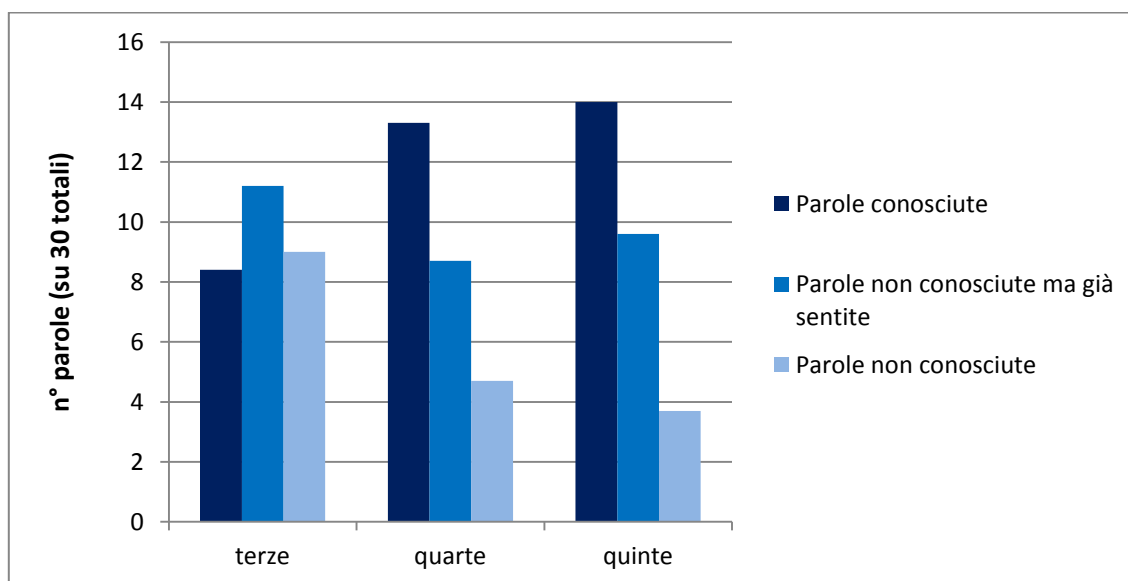


Grafico 4: Valori medi di ogni tipologia di risposta raggruppati per classe.

I bambini di terza hanno risposto soprattutto scegliendo l'opzione "non la conosco ma l'ho già sentita", seguono le risposte "non la conosco" e "la conosco" (che risultano essere la minoranza con il 28% delle risposte). La situazione cambia di molto prendendo in esame le risposte dei bambini di quarta, i quali conoscono la maggior parte delle parole (44%), hanno già sentito il 29% delle parole proposte e non conoscono solo il 16% (4,7 parole). In modo simile si sono espressi anche i

bambini di quinta i quali hanno dichiarato di conoscere poco meno del 50% delle parole testate, di non conoscere solo 3,7 parole (12%) e di non conoscere ma aver già sentito sentito 9,6 parole (32%). Quindi con l'avanzare dell'età le parole sconosciute diventano mano mano più familiari, anche se magari non vengono ancora effettivamente comprese.

Alle percentuali sopra citate vanno aggiunte, per raggiungere il 100% delle risposte per ogni livello di scolarità, quelle (modeste) che si riferiscono alle risposte non date, molto probabilmente per distrazione o dimenticanza, e alle risposte non valide (non è stato registrato uno scarto significativo tra conoscenza dichiarata e conoscenza effettiva). Per considerare una risposta "la conosco" come valida, la definizione corrispondente doveva infatti dimostrare che il bambino aveva un'idea, seppur imprecisa e legata alla sua esperienza personale, del significato di quella parola. In questo modo sono state escluse sia risposte date confondendo la parola con un'altra graficamente simile (ad esempio *senatrice* è stata interpretata da più di un bambino come "macchina per gli agricoltori che semina i semi", cioè la *seminatrice*, oppure *elezioni* è stata definita come "provini", scambiata quindi con *selezioni*, o ancora *costituzione* confusa da alcuni bambini, soprattutto di terza, con *costruzione* dato che la spiegazione fornita è stata: "è quando costruisci un palazzo, un edificio,..."), sia risposte che nascevano da una riflessione errata, come ad esempio quelle date da bambini che credevano che il fascista fosse "una persona che quando ti prendi una grossa storta lui ti fascia" o "è uno che fa le fasce". Anche nel caso di risposte multiple ad un unico quesito la risposta è stata ovviamente scartata.

Considerando le 15 parole del Lessico Elementare emerge che in quinta sono conosciute quasi tutte da almeno un bambino su tre, eccezion fatta per *ministro* che risulta nota a un'esigua minoranza di bambini. Viene dunque confermata l'idea secondo cui a conclusione delle ex scuole elementari, oggi chiamate primarie, la maggior parte dei bambini conosce le parole del Lessico Elementare.

"La conosco" (% sul totale dei bambini per classe)			
	Terza	Quarta	Quinta
cittadino	83,2	92,5	95,9
diritto	38,9	47,5	39,7
legge	55,8	80,0	83,6
sindaco	66,3	82,5	72,6
stato	35,8	46,3	56,2
bandiera	67,4	85,0	89,0
capitale (s.f.)	32,6	55,0	56,2
decisione	71,6	85,0	87,7
ministro	4,2	10,0	5,5
presidente	34,7	61,3	52,1
costituzione	1,1	31,3	47,9
esilio	8,4	31,3	35,6
<i>ideale</i>	36,8	43,8	42,5
indipendenza	8,4	32,5	45,2
sciopero	87,4	82,5	89,0

Tabella 3: Parole del Lessico Elementare conosciute.⁹

Come si vede dalla tabella qui sopra, già in terza più di un bambino su tre conosce 11 delle 15 parole indagate. Delle quattro che risultano sconosciute, due (*ministro* e *indipendenza*) sono state comunque già sentite dalla maggioranza degli intervistati, mentre *costituzione* ed *esilio* risultano essere del tutto sconosciute. In quarta tutte le parole sono conosciute da più del 30% dei bambini, tranne *ministro* che comunque è già stata sentita dal 66,3% dei bambini. In quinta la situazione è analoga: *ministro* continua ad essere poco conosciuta (si registra anzi un peggioramento), mentre *costituzione*, *esilio* e *indipendenza* sono conosciute da una percentuale ancora maggiore di bambini. In quinta ben 9 parole su 15 sono conosciute da oltre il 50% dei partecipanti all'indagine.

Anche le parole del vocabolario fondamentale dovrebbero far parte del patrimonio lessicale di ogni persona che ha completato le scuole elementari (oggi primarie) secondo quanto sostenuto da Tullio De Mauro (cfr. § 1.3.2.). Tuttavia i dati raccolti non confermano completamente questa ipotesi. Se in terza primaria le parole con marca d'uso FO conosciute da meno del 30% dei bambini sono *ministro*, *esilio*, *indipendenza*, *comunista* e *fascista*, in quarta e in quinta sono

⁹ Come nel VdB, anche in questa tabella il grassetto indica l'appartenenza della parola al vocabolario fondamentale, il corsivo a quello di alta disponibilità e il tondo a quello di alto uso.

ancora sconosciute le parole *ministro* (conosciuta dal 5% dei bambini di quinta), *comunista* (nessun bambino la conosce) e *fascista* (conosciuta in quinta dal 19,2% dei bambini e già sentita dal 57,5%). Dunque, se un adulto con licenza elementare molto probabilmente conosce queste tre parole (gli adulti con sola licenza elementare sono tuttavia sempre meno), non sembra possibile affermare che un bambino alla fine del suo percorso all'interno della scuola primaria colga il significato di questi tre vocaboli. In particolare, si può ipotizzare che *comunista* e *fascista* siano parole legate a realtà e fatti storici dei quali difficilmente un bambino di 10 anni sente oggi parlare, considerando anche che i programmi scolastici ministeriali hanno distribuito su tutti i primi otto anni di scuola gli argomenti storici. Tralasciando queste eccezioni, le altre parole del vocabolario fondamentale sono tutte conosciute dalla maggior parte dei bambini che stanno concludendo la scuola primaria, in particolare *cittadino*, *legge*, *sindaco*, *stato*, *bandiera*, *capitale*, *decisione* e *presidente* (8 vocaboli) sono conosciute da più del 50% dei bambini. Tutte queste parole fanno parte anche del Lessico Elementare. Aggiungiamo ora qualche dettaglio per ogni classe scolastica. In terza più del 50% dei maschi conosce le seguenti parole: *cittadino*, *legge*, *sindaco*, *bandiera*, *decisione*, *sciopero*, *inno*, *votazione* (8 vocaboli). Oltre a queste, una bambina su due conosce anche la parola *ideale*. Se invece si prendono in esame le parole conosciute da almeno il 25% degli intervistati, le differenze tra maschi e femmine scompaiono poiché entrambe le categorie conoscono, oltre alle parole già citate (*ideale* compreso), anche *diritto*, *stato*, *capitale* e *presidente* (13 parole, pari al 43% delle parole testate). Maggiori differenze tra i due sessi si hanno invece tra le parole non conosciute da nessuno: se nessun bambino maschio ha infatti dichiarato di conoscere *comunista*, *democrazia*, *propaganda* e *senatrice*, nessuna bambina ha dichiarato di conoscere *costituzione*, *comunista*, *connazionale*, *corruzione*, *democrazia* e *propaganda*. Infine, le parole che in terza sono state dichiarate come non conosciute da almeno due bambini su tre, sia maschi che femmine, sono: *esilio*, *connazionale* e *propaganda*.

In quarta, sono invece i maschi a conoscere una parola in più delle femmine quando si considerano i vocaboli conosciuti da più del 50% degli intervistati: i

maschi conoscono 12 parole (*cittadino, legge, sindaco, stato, bandiera, decisione, presidente, ideale, sciopero, emigrazione, inno, votazione*), le femmine 11 (*cittadino, legge, sindaco, bandiera, capitale, decisione, presidente, sciopero, emigrazione, inno, votazione*). Nonostante qualche piccola differenza tra maschi e femmine rimanga anche prendendo in considerazione le parole conosciute da almeno un bambino su quattro, le parole conosciute da almeno il 25% degli intervistati, sia maschi che femmine, di quarta primaria sono: *cittadino, diritto, legge, sindaco, stato, bandiera, capitale, decisione, presidente, costituzione, esilio, ideale, indipendenza, sciopero, candidata, clandestino, elezione, emigrazione, inno, mafia, parlamento, votazione* (22 parole). Nessun bambino di quarta ha inoltre dichiarato di conoscere la parola *comunista*. Infine *propaganda* e *connazionale* sono le parole con le percentuali più alte di risposte "non la conosco".

Per quanto riguarda le quinte, le parole conosciute da più del 50% dei bambini sono 12, cioè *cittadino, legge, sindaco, stato, bandiera, capitale, decisione, presidente, sciopero, emigrazione, inno, votazione* (i maschi in più conoscono *costituzione*). Se si abbassa la soglia al 25%, invece, le parole conosciute sono 23: alle 12 elencate poco sopra si devono aggiungere *diritto, costituzione, esilio, ideale, indipendenza, candidata, clandestino, democrazia, elezione, mafia* e *senatrice*. Di queste, prendendo maschi e femmine separatamente, *candidata* non compare nell'elenco dei maschi, *democrazia, mafia* e *senatrice* in quello delle femmine. Nella lista delle femmine è però presente la parola *parlamento*, che non raggiunge la soglia del 25% nell'elenco generale (ottenuto facendo la somma delle risposte date dai maschi e quelle date dalle femmine). Le parole con più risposte "non la conosco" sono ancora *propaganda* e *connazionale*, mentre con bassissime percentuali di "la conosco" sono *ministro, comunista, corruzione* e *propaganda*. La parola *propaganda* registra, in altre parole, le seguenti percentuali: 1,4% "la conosco", 57,5% "non la conosco", 38,4% "non la conosco ma l'ho già sentita", 2,7% "risposte non date e non valide".

Dunque le parole che risultano meno conosciute in tutti i livelli di scuola presi in considerazione, dalla terza fino in quinta, sono: *ministro* (che fa parte del Lessico

Elementare e del vocabolario fondamentale), *connazionale*, *propaganda*, *comunista* e *fascista* (questi ultimi due fanno parte del vocabolario fondamentale). Nella tabella e nel grafico che seguono mostriamo infine alcune parole che hanno registrato un particolare incremento con l'aumentare dell'età. La tabella rende più esplicito il dato numerico, mentre il grafico permette di visualizzare rapidamente la "velocità" con cui la singola parola entra nel patrimonio lessicale dei bambini presi in esame.

"La conosco" (% sul totale dei bambini per classe)			
	Terza	Quarta	Quinta
stato	35,8	46,3	56,2
costituzione	1,1	31,3	47,9
indipendenza	8,4	32,5	45,2
democrazia	0	3,8	28,8
emigrazione	15,8	65	86,3
senatrice	1,1	17,5	26

Tabella 4: Parole conosciute con un incremento significativo negli anni.

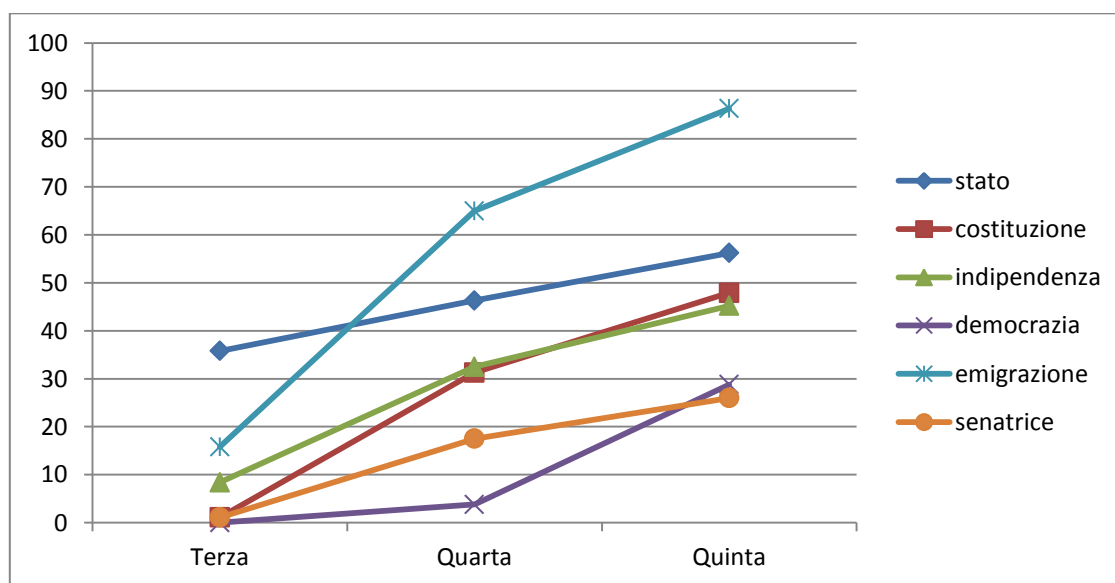


Grafico 5: L'incremento negli anni di sei parole.

3.7. Le definizioni

Come si è già detto, accanto alla risposta chiusa, per ogni parola conosciuta veniva richiesto di spiegare il relativo significato. Alcuni si sono limitati a indicare

uno o più sinonimi, altri hanno provato a fornire una definizione e, tra questi, alcuni hanno aggiunto anche degli esempi per cercare di essere più chiari possibile. È successo anche che alcuni bambini, pur segnando la crocetta su "la conosco" poi hanno scritto "non so spiegare". Questo è significativo per due motivi: sia perché evidenzia la difficoltà del compito assegnato (la riflessione metalinguistica è un processo complesso e che si sviluppa con tempistiche diverse in ogni bambino e comunque molto dopo l'acquisizione linguistica), sia perché mostra come comprendere e usare le parole non implica la capacità di definirle. Quest'ultimo fenomeno si è rivelato in particolare in alcuni questionari in cui il bambino usava una parola per definire un vocabolo conosciuto, ma poi, quando si trattava di definire la parola spontaneamente usata prima, dichiarava di non conoscerla o di averla soltanto già sentita. Tutto ciò va a sostegno di alcuni studi che hanno rilevato come «le competenze lessicali dei bambini sono potenzialmente maggiori, ovvero riconoscono e sanno utilizzare molte più parole di quelle che ritengono di conoscere e che riescono a definire» (Colli Tibaldi *et al.* 2014, p.203).

Definire è un compito impegnativo perché coinvolge la capacità di astrazione, di isolare e selezionare i tratti semantici salienti che distinguono una parola da un'altra. Per questo motivo, le definizioni dei bambini sono spesso ancorate alla concretezza e alla loro esperienza personale. Questo è emerso in modo particolare nelle definizioni date alla parola *sciopero*. Tutti i bambini, infatti, frequentando la scuola, vengono coinvolti indirettamente dallo sciopero: sanno che in caso di sciopero loro non vanno a scuola perché le maestre e le bidelle non ci sono. Ed è il loro punto di vista che appunto traspare dalle loro definizioni:

- *Sciopero significa che per un giorno non c'è scuola;*
- *Saltare per un giorno scuola, lavoro e altre cose che fai di solito;*
- *Lo sciopero è quando le bidelle e i maestri decidono di non venire a scuola;*
- *Sciopero: significa che un posto rimane chiuso fino alla data prevista;*
- *Smettere di lavorare temporaneamente per un qualche fatto;*

Pochi (perlopiù frequentanti la classe quinta) sanno che lo sciopero è in realtà una forma di protesta:

- *Significa protestare e rimanere a casa invece di lavorare;*

- *Sciopero è quando delle persone non vanno al lavoro per protestare contro qualcosa;*
- *Manifestazione lavorativa degli operai di un'azienda o di un istituto;*

Orientate alla concretezza sono anche le spiegazioni date alla parola *stato*, da molti definita come "un insieme di regioni"; tuttavia, guardando tra le definizioni date a questa parola, si possono individuare anche altre strategie definitorie tipiche dei bambini: sono, infatti, usati anche sinonimi ("stato significa nazione" "stato: paese, nazione"), esempi ("un territorio molto grande, Italia, Austria"), generalizzazioni ("un pezzo di terra che sta nel mondo") e riferimenti a conoscenze più specifiche apprese a scuola ("uno stato è una zona di terreno delimitata dove si parla una lingua", "stato è una parte del mondo governato da un presidente").

Nelle definizioni della parola *bandiera* emerge invece un approccio più descrittivo: designando un oggetto, *bandiera* si presta a questo genere di spiegazioni: "la bandiera è un tessuto rettangolare appeso su un bastone in ferro o in legno", "è una stoffa colorata". Non mancano comunque casi in cui, alla descrizione si aggiungono altre informazioni: "è un pezzo di stoffa dove ci sono vari colori con vari significati. Per esempio la bandiera italiana: verde per i prati che abbiamo, rosso per il sangue che abbiamo sudato per averla, bianco per la nostra purezza", "un simbolo d'Italia: bianco, verde, rosso".

Anche definizioni come "le persone scelgono delle cose e a chi piace più la cosa scelta da un altro vota per quello. La maggioranza vince" o "dove ognuno dice la sua e poi si fa la maggioranza" oppure "propria opinione (voto) su qualcosa che si vede", fornite in riferimento alla parola *votazione*, possono essere in un certo senso considerate delle descrizioni: viene spiegato il modo in cui si procede durante una votazione, sia intesa come scelta tra più opzioni, che giudizio su più concorrenti.

Lo stesso si può dire anche per *corruzione*, poiché i pochi che la conoscono ripropongono nella spiegazione una scena per loro esemplare: "vuol dire tipo: che un prigioniero ti dà dei diamanti e in cambio devi dargli la libertà cioè lasciarli rapinare le banche" oppure "è quando un giudice, ad esempio, non fa andare in prigione una persona cattiva, per soldi".

Oltre alle modalità già esemplificate, ci sono stati anche casi di ricostruzione del significato facendo affidamento alla forma della parola presentata. Nell'arduo compito di spiegare che cosa significa *decisione*, ad esempio, molti hanno cercato dei sinonimi ("scelta", "è una promessa"), ma molti hanno anche ricondotto la parola al verbo dal quale deriva, dando spiegazioni del tipo "è quando decidi di fare qualcosa", "decidere una cosa e rispettarla".

La riflessione sulla forma morfologica della parola è evidente anche in alcune definizioni (sbagliate) date a *fascista* (cfr. § 3.6.) e *comunista*, parole generalmente poco o affatto conosciute. Ne sono un esempio per quanto riguarda *comunista* definizioni come "una persona che lavora in comune" o "una persona del comune". È interessante notare come il suffisso -ista suggerisca a tutti che la parola (sconosciuta) indica una persona. Nessuno ha infatti ricondotto la parola a referenti inanimati, neanche coloro che hanno proposto congetture diverse come "quando uno ha fatto la comunione" o "persona che crede in dio".

Una menzione meritano sicuramente anche le definizioni date a *ideale*, che si dividono fondamentalmente in due tipi: da una parte quelle che citano idea come sinonimo ("è un'idea", "significa idea di una cosa"), dall'altra quelle che rinviano alla perfezione ("ideale: cosa perfetta", "qualcosa che è perfetto"). Nessuno tuttavia, nonostante il contesto frasale nel quale la parola è stata presentata,¹⁰ ha inteso *ideale* come forma di aspirazione con nobili scopi, che evidentemente non è un'accezione ancora conosciuta.

Come si può notare anche nelle definizioni già riportate, nelle spiegazioni date a molte parole si ritrovano espressioni generiche tipiche del linguaggio dei bambini, come *cosa*, *persona*, *uno*, *quello*, etc. Solo per fare qualche esempio: "[il connazionale è] uno che sta nella stessa nostra nazione", "candidata: una persona che si è proposta", "la legge è una cosa da rispettare". Tipico è anche l'uso di "è quando" come forma introduttiva alla spiegazione: "[indipendenza] è quando non sei sotto il controllo di altri stati", "[esilio] è quando viene cacciato dal proprio Paese".

Per una selezione delle definizioni di ogni vocabolo si rimanda all'appendice 3.

¹⁰ "Gandhi è famoso per il suo ideale di pace."

3.8. Riflessioni conclusive

Dall'indagine condotta emergono perlopiù conferme delle premesse teoriche da cui siamo partiti. Innanzitutto è stato verificato il carattere incrementale del lessico: nei tre livelli scolastici presi in considerazione si è assistito, infatti, ad un aumento del numero di parole conosciute e a una diminuzione di quelle sconosciute. Questo conferma che all'aumentare dell'età aumentano anche le competenze lessicali dei bambini, e in particolare si è registrato uno stacco maggiore tra la terza e la quarta primaria. Tuttavia, ribadiamo che quanto rilevato dallo studio va considerato come una tendenza non generalizzabile poiché la ricerca è stata condotta per un numero limitato di parole e i bambini testati, per quanto numerosi, appartengono ad un'unica area geografica.

Viene confermata inoltre la conoscenza da parte della maggior parte dei bambini delle parole che fanno parte del Lessico Elementare (fa eccezione solo *ministro*). Dunque, nonostante siano passati ormai vent'anni dalla data di pubblicazione del *Lessico Elementare* e nonostante le trasformazioni sociali avvenute nel frattempo, le parole conosciute dai bambini della scuola primaria, sono grossomodo ancora le stesse.

Come ci si aspettava inoltre, alla fine dei primi cinque anni di scuola, il VdB non è ancora completamente entrato nel patrimonio lessicale dei bambini, anche se dai dati risulta che un bambino su quattro conosce due terzi delle parole presentate.

Non confermano del tutto le aspettative invece le parole del vocabolario fondamentale, che proprio a conclusione delle scuole primarie dovrebbe esser completamente conosciuto. Dall'indagine è infatti risultato che tre parole con marca FO sono poco conosciute: *ministro*, *comunista* e *fascista*. La prima, pur essendo conosciuta da un'esigua minoranza di intervistati è comunque stata sentita dal 75% dei bambini, *fascista* non è molto conosciuta ma è già stata sentita dai più, *comunista* non è invece conosciuta da nessuno. Si può ipotizzare che per quanto riguarda *fascista* e *comunista* il motivo di un così scarso livello di conoscenza sia dovuto al riferirsi a concetti storici molto precisi e complessi, di cui si ha poche occasioni di parlare, soprattutto con i bambini. Se a cavallo tra gli

anni Settanta e gli anni Ottanta (quando è stato realizzato il VdB) queste parole potevano considerarsi patrimonio lessicale di tutti coloro che avevano concluso la scuola elementare, tanto da venir inserite tra le 2000 del vocabolario fondamentale dell'italiano, oggi non sembra possibile dire lo stesso, almeno per quanto riguarda i bambini in uscita dalla scuola primaria.

Per quanto riguarda le definizioni, infine, si può affermare che quelle raccolte con i questionari presentano le caratteristiche che sono state individuate come peculiari delle definizioni dei bambini: la tendenza alla concretezza, il forte legame con le proprie esperienze personali, le generalizzazioni, l'uso di espressioni generiche come "è quando" "è uno", l'uso di descrizioni, etc.

Nel complesso dunque le premesse teoriche da cui è partita la ricerca trovano conferma nei dati ottenuti.

Conclusioni

Far parte di una società significa avere a che fare con le altre persone che costituiscono quella società, significa creare collegamenti e condivisioni con loro per poter sfruttare appieno le enormi potenzialità che le persone messe assieme hanno. E affinché ognuno possa dare il proprio contributo alla costruzione di un bene comune, è necessario innanzitutto condividere un codice linguistico con cui comunicare, è necessario che chi parla sappia scegliere le parole che meglio rispecchiano il concetto che vuole esprimere e, sull'altro versante, è necessario che chi ascolta sappia decifrare il messaggio che riceve attribuendo ad ogni parola la corrispettiva sfumatura di significato che quel vocabolo assume in quel preciso contesto. Solo così, con una corrispondenza quasi perfetta tra ciò che viene detto e ciò che viene recepito, ci saranno le basi per la costruzione di una sapere comune.

Diventa perciò assolutamente necessario assicurarsi che ogni cittadino (nel nostro caso italiano) conosca e sia in grado di usare il nucleo di parole che formano la struttura portante del 98% degli enunciati che vengono formulati: le 7000 parole del vocabolario di base. È innanzitutto a questa fascia di lessico, e in particolare alle parole fondamentali, che va dunque dedicata la massima attenzione, anche e soprattutto nel costruirsi dei progetti di apprendimento. Se, infatti, molti studi mostrano una correlazione tra competenze lessicali e contesto socio-culturale e familiare, non va trascurato il ruolo che la scuola può e deve rivestire nello sviluppo del vocabolario di ognuno. Non solo gli insegnanti di italiano devono porre tra i loro obiettivi il lessico, ma anche quelli che insegnano le altre discipline, perché ogni argomento offre nuove parole, o nuove accezioni di parole già conosciute, con cui familiarizzare. Conoscere le parole del VdB è importante non solo per il significato di cui ogni parola è portatrice, ma anche perché per spiegare il significato di parole nuove si ricorre spesso alle parole più facilmente conosciute dal parlante e dall'intera comunità: le parole contenute nel VdB svolgono un ruolo di primo piano anche durante l'espansione del lessico, sono i mattoncini su cui si fondano le altre conoscenze lessicali. In conclusione si può dire che possedere il VdB consente di partecipare alla conoscenza e alla vita insieme.

Dalle considerazioni fatte in questo lavoro emerge il ruolo di grande rilievo che il lessico riveste all'interno della lingua ed è importante trovare dei metodi di misurazione delle conoscenze lessicali dei parlanti sia per intervenire là dove ci sono delle lacune sia per approfondire e affinare la conoscenza di quei vocaboli con i quali si ha già familiarità. Qualità e quantità sono infatti due aspetti che vanno affrontati parallelamente 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» (Ferrerri 2005, p.38). Tuttavia dalla letteratura in merito risulta che non ci sono ad oggi dei test standardizzati per valutare le competenze lessicali e questa è una grave mancanza poiché nega la possibilità di fare confronti con situazioni passate o con realtà geograficamente e socialmente diverse.

Non sottovalutare le potenzialità, non solo linguistiche, ma anche sociali, delle conoscenze lessicali e ricercare strategie di valutazione standardizzate sono dunque i propositi che con maggior urgenza emergono dal presente lavoro.

Bibliografia

- LUIGI APRILE, *L'evoluzione e la rappresentazione del lessico nell'infanzia*, «Età evolutiva», 40 (1991), pp.120-127.
- GRAZIA BASILE, (2012), *La conquista delle parole. Per una storia naturale della denominazione*, Roma, Carocci.
- GRAZIA BIORCI, LUCIA FERLINO, MICAELA ROSSI, (2006), *Imparare dai bambini: riflessioni a margine di un'esperienza*, in *Linguaggio, mente, parole. Dall'infanzia all'adolescenza*, a cura di Immacolata Tempesta e Maria Maggio, Milano, Franco Angeli, pp. 147-157.
- VALENTINA BISCONTI, *La svolta lessicografica di Tullio De Mauro e i dizionari contemporanei*, «Chroniques italiennes web», 23 (2/2012).
Scaricabile al seguente indirizzo: <http://chroniquesitaliennes.univ-paris3.fr/PDF/Web23/Bisconti-La-svolta-lessicografica.pdf>
- ANTONIETTA BISETTO, SERGIO SCALISE, (2008), *La struttura delle parole*, Bologna, Il Mulino.
- UMBERTA BORTOLINI, CARLO TAGLIAVINI, ANTONIO ZAMPOLLI, (1971), *Lessico di frequenza della lingua italiana contemporanea*, Milano, IBM Italia.
Introduzione scaricabile al seguente indirizzo:
http://www.ilc.cnr.it/AZ_bibliography/Z018.PDF
- FILIPPO BOSCHI, LUIGI APRILE, IDA SCIBETTA, (1992), *Le parole e la mente. Studi e ricerche sullo sviluppo della competenza lessicale nel ragazzo*, Firenze, Giunti.
- FEDERICA CASADEI, (2003), *Lessico e semantica*, Roma, Carocci.

- CHIARA COLLI TIBALDI, LAURA DELUIGI, MICHELA FRAIRE, (2014), *Le competenze lessicali nella scuola primaria: un'indagine sul lessico sensoriale*, in *L'italiano per capire* a cura di Adriano Colombo e Gabriele Pallotti, Roma, Aracne, pp.195-207.
- EMILIO D'AGOSTINO, *Il lessico di frequenza dell'italiano parlato e la didattica dell'italiano*, «Quaderns d'Italià», 3 (1998), pp. 9-28.
- ALAN DAVIES, (2008), *Che cosa ci dicono i test lessicali riguardo alla padronanza linguistica?*, in *Lessico e apprendimenti. Il ruolo del lessico nella linguistica educativa*, a cura di Monica Barni, Donatella Troncarelli e Carla Bagna, Milano, Franco Angeli, pp. 99-110.
- SALVATORE DE MASI E MARIA MAGGIO, (2006), *Le parole dei bambini al termine della Scuola Primaria*, in *Linguaggio, mente, parole. Dall'infanzia all'adolescenza*, a cura di Immacolata Tempesta e Maria Maggio, Milano, Franco Angeli, pp.174-183.
- TULLIO DE MAURO, (2003), *Guida all'uso delle parole. Parlare e scrivere semplice e preciso per capire e farsi capire*, XII ed., Roma, Editori Riuniti.
- TULLIO DE MAURO, (2008), *Parole come semi*, in *Lessico e apprendimenti. Il ruolo del lessico nella linguistica educativa* a cura di Monica Barni, Donatella Troncarelli e Carla Bagna, Milano, Franco Angeli, pp.27-46.
- GIANFRANCO DENES, (2009), *Parlare con la testa. Le basi neurologiche e la struttura del linguaggio*, Bologna, Zanichelli.
- SILVANA FERRERI, (2005), *L'alfabetizzazione lessicale. Studi di linguistica educativa*, Roma, Aracne.

- SILVANA FERRERI, (2006), *Parole tra quantità e qualità*, in *Linguaggio, mente, parole. Dall'infanzia all'adolescenza*, a cura di Immacolata Tempesta e Maria Maggio, Milano, Franco Angeli, pp. 131-146.
- STEFANO GENSINI, MASSIMO VEDOVELLI, (1983), *Strumenti per la programmazione dell'educazione linguistica: il glotto-kit*, in *Teoria e pratica del glotto-kit. Una carta d'identità per l'educazione linguistica* a cura di Stefano Gensini e Massimo Vedovelli, Milano, Franco Angeli.
- GISCEL CALABRIA, (2008), *La conoscenza del vocabolario comune negli studenti dalle elementari alla maturità*, in *Lessico e apprendimenti. Il ruolo del lessico nella linguistica educativa* a cura di Monica Barni, Donatella Troncarelli e Carla Bagna, Milano, Franco Angeli, pp. 187-201.
- GISCEL SICILIA, (2014), *Le competenze lessicali tra autovalutazione e uso: un'indagine nelle scuole di Palermo*, in *L'italiano per capire* a cura di Adriano Colombo e Gabriele Pallotti, Roma, Aracne, pp.169-180.
- ELISABETTA JEŽEK, (2005), *Lessico. Classi di parole, strutture, combinazioni*, Bologna, Il Mulino.
- KENDALL A. KING, ALISON MACKEY, (2008), *L'acquisizione linguistica*, Bologna, Il Mulino.
- MARIA GIUSEPPA LO DUCA, *Le definizioni dei bambini*, «Italiano e Oltre», I (1986), pp. 124-127.
- L. MARCONI, M. OTT, E. PESENTI, D. RATTI, M. TAVELLA, (1994), *Lessico elementare. Dati statistici sull'italiano scritto e letto dai bambini delle elementari*, Bologna, Zanichelli.

- MIUR, (2012), *Indicazioni nazionali per il curricolo della scuola dell'infanzia e del primo ciclo d'istruzione*.
- ELENA PAPA, (2014), «*Il deserto è quando...*». *Competenze definitorie e competenze disciplinari nella scuola primaria*, in *L'italiano per capire* a cura di Adriano Colombo e Gabriele Pallotti, Roma, Aracne, pp.181-193.
- MARIA CHIARA STARACE, (1983), *Sondaggi sulla conoscenza del vocabolario di base nella scuola media*, in *Teoria e pratica del glotto-kit. Una carta d'identità per l'educazione linguistica* a cura di Stefano Gensini e Massimo Vedovelli, Milano, Franco Angeli.
- ANNA M. THORNTON, CLAUDIO IACOBINI, CRISTINA BURANI, (1997), *BDVDB. Una base di dati sul vocabolario di base della lingua italiana*, Roma, Bulzoni Editori.
- MASSIMO VEDOVELLI, (1983), *Parametri di conoscenza ricettiva di vocabolario di base*, in *Teoria e pratica del glotto-kit. Una carta d'identità per l'educazione linguistica* a cura di Stefano Gensini e Massimo Vedovelli, Milano, Franco Angeli.

Appendice 1

Lista delle parole testate.

	Rango			
1	cittadino	1539	vocabolario fondamentale	14
2	diritto	1923	vocabolario ad alto uso	9
3	legge	1076	<i>vocabolario ad alta</i>	7
4	sindaco	1673	<i>disponibilità</i>	
5	stato	768	Legenda dei caratteri utilizzati.	
6	bandiera	2511		
7	capitale (s.f.)	2672		
8	decisione	2105		
9	ministro	3750		
10	presidente	2448		
11	costituzione	4817		
12	esilio	4482		
13	<i>ideale</i>	5592		
14	indipendenza	4606		
15	sciopero	4505		
16	<i>candidata</i>			
17	clandestino			
18	comunista			
19	connazionale			
20	<i>corruzione</i>			
21	democrazia			
22	elezione			
23	<i>emigrazione</i>			
24	fascista			
25	inno			
26	<i>mafia</i>			
27	parlamento			
28	propaganda			
29	<i>senatrice</i>			
30	<i>votazione</i>			

**Tabella 5: Lista delle parole testate.
Rango di quelle che appartengono al
Lessico Elementare, oltre che al VdB.**

*Ho preso una **decisione**: da domani mangerò meno dolci.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*Tutti i **cittadini** hanno sia diritti che doveri.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*Il **sindaco** di Turriaco è Enrico Bullian.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*La **costituzione** italiana è nata nel 1948.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*La **legge** va rispettata da tutti.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*Dire la propria opinione è un **diritto**.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*Napoleone è stato costretto all'**esilio**.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*L'uomo che hanno mostrato alla tv era un **clandestino**.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*Nel 1947 l'India ottenne l'**indipendenza**.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*Prima delle partite dei mondiali ogni squadra canta il proprio **inno**.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*Per scegliere il nuovo caposquadra faremo una **votazione**.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*Barack Obama è il primo **presidente** degli Stati Uniti di origine afroamericana.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*Il **parlamento** italiano è formato da due camere, quella dei senatori e quella dei deputati.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*Sull'**atlante** ogni **stato** ha un colore diverso.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....
.....

*Il vincitore della maratona è un nostro **connazionale**!*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....

.....

*L'**emigrazione** ha interessato per molto tempo l'Italia.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....

.....

*Mia zia, che abita a Roma, conosce personalmente una **senatrice**.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....

.....

*Il **ministro** dell'agricoltura la prossima settimana visiterà le cantine di Cormons.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....

.....

*Secondo alcuni l'uomo della foto era un **fascista**, secondo altri no.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....

.....

*Durante il periodo di **propaganda** la città si riempie di locandine e manifesti.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....

.....

*La nuova **candidata** del partito XXL ha fatto un discorso alla radio.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....

.....

*Il bisnonno di Piero era forse **comunista**?*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....

.....

*Combattere la **corruzione** non è una cosa facile.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....

.....

*La **democrazia**, anche se diversa da oggi, esisteva già nell'antichità.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....

.....

*Tutti si chiedono chi vincerà le prossime **elezioni**.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....

.....

*Purtroppo all'estero l'Italia è spesso associata alla pizza e alla **mafia**.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....

.....

*Gandhi è famoso per il suo **ideale** di pace.*

- La conosco Non la conosco Non la conosco ma l'ho già sentita

Spiega con parole tue il significato della parola

.....

.....

Appendice 3

Esempi di definizioni date per ogni parola.

	Definizione	Classe	Valida
Bandiera	Una bandiera è un tessuto colorato con i colori delle nazione	III	Sì
	È tipo uno striscione di stoffa con il disegno della regione	IV	Sì
	Logo o stemma di una nazione, di una regione, di un continente,...	V	Sì
Candidata	Candidata è qualcuno che si offre	III	Sì
	Candidata vuol dire persona che vuole essere eletta	IV	Sì
	Candidata: una persona che si è proposta	V	Sì
Capitale	È una città molto importante	III	Sì
	La capitale è il posto più importante di tutto il paese	IV	Sì
	Città a capo dello stato, dove ci sono i ministri	V	Sì
Cittadino	I cittadini sono le persone che abitano in una città	III	Sì
	I cittadini sono gli abitanti di una città	IV	Sì
	Cittadini vuol dire residenti con cittadinanza italiana	V	Sì
Clandestino	Clandestino: persona proveniente dall’Africa	III	Sì
	È una persona che arriva da un altro paese senza però avere i documenti	IV	Sì
	Significa che non era di quel paese ed era venuto in quel paese illegalmente	V	Sì
Comunista	Una persona che lavora in comune	III	No
	Quando uno ha fatto la comunione	IV	No
	Una persona del comune	V	No
Connazionale	Significa che abita nel nostro stato	III	Sì
	Uno che sta nella stessa nostra nazione	IV	Sì
	Un connazionale è una persona che appartiene alla stessa nazione	V	Sì
Corruzione	Vuol dire tipo: che un prigioniero ti da dei diamanti e in cambio devi dargli la libertà cioè lasciarli rapinare le banche	III	Sì
	È quando un giudice, ad esempio, non fa andare in prigione una persona cattiva, per soldi	IV	Sì
	Quando uno imbrogli qualche persona	V	No
Costituzione	È quando costruisci un palazzo, un edificio,...	III	No
	È formata da tanti articoli	IV	Sì
	Legge più importante per lo stato italiano	V	Sì

Decisione	Decisione vuol dire che hai deciso di fare qualche cosa	III	Sì
	È quando decidi di fare qualcosa	IV	Sì
	Decidere una cosa e rispettarla	V	Sì
Democrazia	Democrazia: stato governato da una famiglia reale	IV	No
	Vuol dire che è il popolo che decide	V	Sì
Diritto	Diritto significa una cosa che tu vuoi fare e si può anche fare	III	Sì
	Diritto: qualcosa che puoi fare	IV	Sì
	Poter fare una cosa	V	Sì
Elezione	Si vota per il nuovo sindaco, il nuovo presidente	III	Sì
	Delle persone che scelgono tra i candidati chi potrebbe fare il sindaco o qualcos'altro	IV	Sì
	È un "concorso" dove i cittadini votano chi sarà il prossimo presidente	V	Sì
Emigrazione	Emigrazione è quando qualcuno si sposta	III	Sì
	Quando qualcuno se ne va dal proprio paese	IV	Sì
	Andare via dal proprio paese per alcuni motivi	V	Sì
Esilio	Esilio: morire	III	No
	È quando viene cacciato dal proprio Paese	IV	Sì
	Esilio vuol dire essere mandato via in un altro paese perché deve aver fatto qualcosa di male	V	Sì
Fascista	È uno che fa le fasce	III	No
	Persona che vuole essere a capo di tutti e non accetta quelli diversi da lui	IV	Sì
	Persone italiane che seguivano Mussolini	V	Sì
Ideale	Ideale significa qualcosa che è un'idea molto buona	III	Sì
	Ideale: cosa perfetta	IV	Sì
	È un'idea	V	Sì
Indipendenza	Fare le cose da solo	III	Sì
	L'indipendenza è essere autonomi	IV	Sì
	Uno stato non dipende dall'altro	V	Sì
Inno	È una canzone	III	Sì
	È una canzone che si canta alle partite	IV	Sì
	Canzone dedicata o fatta per uno stato, squadra	V	Sì
Legge	È una regola che bisogna rispettare	III	Sì
	Sono delle regole che vanno eseguite	IV	Sì
	Un insieme di regole da rispettare	V	Sì
Mafia	La mafia è tanta gente che ammazza altra gente	III	Sì
	Gruppo di persone mal intenzionate che certe volte uccidono persone innocenti	IV	Si

	È la malavita che fa atti clandestini e non di aiuto per il popolo	V	Sì
Ministro	Ad esempio Matteo Renzi è un ministro	III	Sì
	Persona che decide una cosa per tutti	IV	Sì
	Ministro chi si occupa di un argomento specifico	V	Sì
Parlamento	Sono le persone che decidono le leggi	III	Sì
	Posto dove i politici si raggruppano per parlare e per discutere	IV	Sì
	Aula in cui i politici decidono leggi	V	Sì
Presidente	È una persona come il sindaco solo che si occupa dello stato	III	Sì
	Una persona che sta a capo di uno stato	IV	Sì
	Comanda, con l'aiuto di altre persone, uno stato	V	Sì
Propaganda	"Festa"	III	No
	Festa dove si vendono cose nei mercatini	IV	No
	Pubblicità	V	Sì
Sciopero	Sciopero significa che per un giorno non c'è scuola	III	Sì
	Lo sciopero è quando le bidelle e i maestri decidono di non venire a scuola	IV	Sì
	Sciopero è quando delle persone non vanno al lavoro per protestare contro qualcosa	V	Sì
Senatrice	Qualcuno che comanda su un posto	III	No
	Una signora con tanto seno (quello che hanno le donne)	IV	No
	Persona "donna" che lavora al senato	V	Sì
Sindaco	Il capo di un paese o una città	III	Sì
	È la persona che governa la città	IV	Sì
	Una persona che è a capo di un paese	V	Sì
Stato	Una nazione	III	Sì
	Stato è un insieme di regioni	IV	Sì
	Stato è una parte del mondo governato da un presidente	V	Sì
Votazione	Significa votare	III	Sì
	Dove ognuno dice la sua e poi si fa la maggioranza	IV	Sì
	Una decisione che si prende in gruppo	V	Sì

Ringraziamenti

Un primo sentito grazie va alla dirigente scolastica dell'Istituto Comprensivo "Dante Alighieri" di San Canzian d'Isonzo, la professoressa Annamaria Bonato, e alla sua collaboratrice, la dottoressa Alessandra Rea: senza la loro collaborazione non sarebbe stato possibile ottenere i dati necessari per la ricerca.

Un grazie speciale va anche a tutti gli insegnanti che con grande generosità mi hanno concesso il loro prezioso tempo per somministrare i questionari: sono stati tutti molto disponibili e collaborativi.

Grazie anche a tutte le segretarie dell'Istituto che hanno pazientemente risposto alle mie numerose telefonate.

Devo inoltre ringraziare le professoresse Maria Teresa Vigolo e Loredana Corrà che mi hanno dato il sostegno necessario per realizzare questo lavoro.

Un grande grazie va anche alle mie coinquiline Giulia e Silvia che mi hanno supportata nelle prime e più complicate fasi del lavoro.

Infine ringrazio la mia famiglia, che mi ha permesso di fare quest'esperienza padovana.