

LE PAROLE DEL RAV

L'analisi testuale per la metavalutazione

Analisi dei punti di forza e di debolezza:
area del Successo scolastico





INVALSI - Via Ippolito Nievo 35, Via Marcora 18/20, 00153 Roma.

Le parole del Rav. L'analisi testuale per la metavalutazione

A cura di:

Donatella Poliandri, Ughetta Favazzi, Monica Perazzolo, Isabella Quadrelli, Emanuela Vinci

Emanuela Vinci ha elaborato i dati e redatto il presente capitolo.

Alla riflessione sull'autovalutazione e valutazione delle scuole, e al progetto PON Valu.E, hanno contribuito in questi anni: Fabio Alivernini, Mattia Baglieri (consulente), Paola Bianco, Roberta Cristallo (consulente), Nicoletta Di Bello, Graziana Epifani, Stefano Famiglietti (responsabile Settore Web), Ughetta Favazzi, Brunella Fiore (assegnista), Francesca Fortini, Michela Freddano, Letizia Giampietro, Filippo Gomez Paloma (consulente), Angela Litteri, Beba Molinari (consulente), Lorenzo Mancini, Sara Manganelli, Daniela Marinelli, Flora Morelli, Enrico Nerli Ballati, Monica Perazzolo, Donatella Poliandri (responsabile dell'Area di Ricerca INVALSI - Innovazione e Sviluppo e del progetto PON Valu.E), Elisabetta Pranterà, Isabella Quadrelli (consulente), Maria Ranieri (consulente), Sara Romiti, Simone Russo, Stefania Sette, Consuelo Torelli (assegnista), Emanuela Vinci.

La dimensione del **Successo Scolastico** afferisce all'Area degli Esiti presente nel RAV e definisce gli esiti degli studenti nel breve e nel medio periodo. È importante per le scuole sostenere il percorso scolastico degli studenti, prestando particolare attenzione agli studenti in ritardo, con debiti formativi, promossi con il minimo dei voti, e riducendo al minimo l'incidenza dei trasferimenti e degli abbandoni. Il criterio di qualità indicato nella Rubrica di valutazione specifica che "La scuola garantisce il successo formativo degli studenti". Le domande guida orientano le scuole nell'individuazione sia degli elementi di forza sia di quelli di debolezza che si riferiscono, in particolare, agli studenti non ammessi; alla loro distribuzione per fascia di voto conseguito all'Esame di stato; all'entità degli abbandoni e trasferimenti.

2.1 Introduzione

Il lavoro relativo all'analisi dei testi dei campi aperti riguardanti i punti di forza e di debolezza, descritti dalle scuole nel proprio percorso di valutazione, si colloca all'interno del processo, in parte già avviato, della Metavalutazione del Rapporto di Autovalutazione (RAV). Uno degli obiettivi espliciti di tale studio è, a partire dal processo riflessivo condotto dalla scuola durante l'autovalutazione così come cristallizzato nei testi redatti dalla scuola stessa, comprendere quale sia la reale efficacia degli strumenti elaborati in termini di supporto e di utilità rispetto all'intero percorso.

Nello specifico, cerchiamo di verificare il livello di coerenza tra punti di forza e di debolezza, individuati dalle scuole, con i giudizi espressi nelle rubriche di valutazione e la scelta delle priorità. È opportuno ricordare che la riflessione da parte delle scuole su tali aspetti è sollecitata dalle domande guida presenti alla fine di ciascuna sezione; pertanto possiamo pensare che, nella elaborazione dei testi, esse abbiano ripreso e rielaborato tematiche e parole prossime, in termini di significato, a quelle proposte nel RAV.

L'analisi del profilo lessicale prende l'avvio dall'analisi dei testi inseriti dalle scuole nei punti di forza e nei punti di debolezza della dimensione del "Successo

scolastico" che si compone di due aspetti definiti da altrettanti indicatori, ossia gli *Esiti degli scrutini* e i *Trasferimenti e abbandoni*. Nelle domande guida a essi relativi, i temi proposti vertono sulla distribuzione degli studenti in determinate situazioni positive (ammissioni alle classi successive), o in situazioni negative (abbandoni, sospensioni, non ammissioni alle classi successive).

In questa sezione del RAV, rileviamo elementi oggettivi della realtà scolastica che le scuole acquisiscono direttamente dal Ministero della Pubblica Istruzione, Università e Ricerca (MIUR) - o da altre fonti ufficiali-; ciò fa ritenere che, ancor più che per altre aree del RAV, le scuole abbiano utilizzato una terminologia aderente a quanto suggerito nelle domande guida, facendo riferimento ai dati presenti e argomentando in maniera più dettagliata in caso di situazioni particolari che richiedono una descrizione più dettagliata o di particolari condizioni da esporre in maniera più analitica.

2.2 Metodo: l'analisi lessicale sui punti di forza e di debolezza dell'area Successo scolastico

Considerando i campi aperti in esame come frammenti di testo indipendenti, abbiamo impostato alcune analisi preliminari per una prima esplorazione lessicale del corpus testuale¹. Tali campi aperti si caratterizzano come corpus di medie dimensioni, che possono essere considerati analogamente alle risposte a domande aperte di un questionario e che rappresentano le nostre unità di analisi.

Entrando nel merito dei punti di forza, ci troviamo ad esplorare un corpus costituito da circa 81.912 occorrenze e da quasi 5.000 differenti forme grafiche (4.830 *type*), rappresentato da 725 RAV compilati dalle scuole, di cui il 33% del II ciclo, e il 67% del I ciclo di istruzione.

Tale corpus rappresenta la percentuale totale di copertura delle parole, non essendo stata inserita la *stop list*². Essendo un corpus di medie dimensioni è stato possibile calcolare il quoziente *type/token* (cfr. Glossario) che è pari a 0,059, valore che è al di sotto del valore soglia del 20% indicato da Bolasco (2013) per considerare il corpus come adeguato per procedere a una successiva analisi testuale. Tale coefficiente consente di conoscere l'eterogeneità delle parole utilizzate, dando una prima informazione sulla ricercatezza del linguaggio utilizzato dalle scuole per descrivere la loro situazione in termini positivi e negativi: più il corpus del testo è grande, minore è il valore che ci aspettiamo di individuare.

Considerando invece i campi aperti relativi ai punti di debolezza dell'area del

successo scolastico ci troviamo a esaminare 719 RAV compilati.

Le analisi dei punti di debolezza redatti dalle scuole mostrano i seguenti valori: il numero delle occorrenze risulta ridotto a circa 56.560; tuttavia il numero delle forme grafiche risulta essere più o meno analogo al precedente (4.664 *type*).

Relativamente a questo frammento di testo il quoziente *type/token* rilevato è pari a 0,082. Ciò lascia supporre che non vi siano evidenti differenze di utilizzo delle forme grafiche nel descrivere situazioni positive (punti di forza) e situazioni negative (punti di debolezza) da parte delle scuole; tuttavia il numero delle parole utilizzate è certamente inferiore nel secondo caso.

Tale tendenza è stata riscontrata anche in altre analisi analoghe implementate su differenti aree di contenuto. Questo atteggiamento diffuso vede le scuole descrivere in maniera più estesa e dettagliata i punti di forza rispetto ai punti di debolezza; questi ultimi sono presentati sempre in maniera sintetica, lasciando presupporre il fine - più o meno esplicito - di minimizzarli 'parlandone' di meno. Questo fenomeno, come sarà possibile leggere in tutto il documento da noi redatto, è presente in tutte le aree analizzate, come già accennato, rappresentando una delle evidenze empiriche che emerge da questo genere di analisi. Ciò ci fa riflettere anche dell'importanza di una diffusione della cultura della valutazione come volta al miglioramento, e non intesa come verifica e controllo al fine di elargire premi e punizioni.

La valutazione intesa come formativa, come rilevante anche per coloro

¹ Il termine corpus indica l'insieme di tutti i campi aperti presenti nello strumento (RAV); i campi aperti relativi ai singoli punti di forza e punti di debolezza rappresentano le "unità di contesto o frammenti" (Bolasco, 2013).

² Si tratta di una lista di esclusione che consente di escludere dall'analisi testuale alcune forme grafiche

presenti nei testi, ma ritenute poco significative ai fini delle analisi, come ad esempio congiunzioni, preposizioni, articoli.

Il software QDAMiner WordStat permette di definire una *stop list* personalizzata: le parole in essa contenute non vengono inserite nel computo delle analisi implementate. (Cfr. Glossario)

che la conducono è ancora un approccio ancora poco diffuso nel mondo della scuola; questo può rendere – non impossibile - ma più difficile da un punto di vista empirico un'esposizione maggiormente chiara, argomentata e trasparente dei propri punti di debolezza.

La fase successiva del lavoro si è incentrata sull'analisi delle liste di parole dei punti di forza, attraverso la ripartizione delle frequenze delle parole in fasce (bassa fino a 78 occorrenze, media, e alta da 229 occorrenze in poi, secondo le indicazioni di Bolasco, 2013)³.

Nella fascia con frequenze più elevate, (tabella 1.1 in Appendice)⁴ si concentrano le parole cosiddette 'vuote' dei

testi: congiunzioni, preposizioni, articoli, ecc.⁵ Le prime *parole piene o principali*⁶ (Bolasco, 1999), che si evidenziano nella figura 1.1, si riferiscono in primo luogo ai soggetti coinvolti nel processo di valutazione: la parola "scuola" è citata 1.111 volte, seguita da "studenti" (1.046), "alunni" (800) e "classe" con 746 occorrenze.

Tra le parole successive si collocano quelle che indicano delle tendenze in atto rispetto ai temi legati al successo scolastico; si tratta delle parole "ammessi" (592 occorrenze), "percentuale" (585), "media" (544), "successiva" 487, "valutazione" 480, che le scuole hanno utilizzato per operare un confronto tra la loro situazione e i riferimenti territoriali presenti nel RAV.

Figura 2.1 - Parole piene a frequenza elevata relative ai punti di forza dell'area Successo scolastico



³ Partendo dal fondo della distribuzione di frequenze, ovvero dagli hapax, la fascia bassa di frequenza termina quando compare il primo salto nella sequenza delle frequenze (appare un valore mancante nella sequenza dei numeri di occorrenze); la fascia media termina e inizia quella alta quando si incontra l'ultima parola con la stessa frequenza di quella precedente (partendo dall'alto si individua la prima frequenza che si ripete due volte).

⁴ In appendice al presente capitolo sono state riportate le tabelle con tutte le parole presenti nelle fasce di frequenza; le nuvole di parole presentate nei

testi rappresentano le parole piene estrapolate dalle medesime fasce di frequenza a seguito dell'inserimento della *stop list*.

⁵ Nella fascia a bassa frequenza ci sono anche gli hapax, ossia i termini che compaiono nel testo una sola volta: nei punti di forza essi rappresentano il 46,8% della partizione.

⁶ Nelle tre fasce di frequenza si articolano progressivamente i differenti tipi di parole: parole vuote, parole strumentali, parole principali (Bolasco, 1999).

L'obiettivo fin qui raggiunto, ossia di esplorare i testi contenuti nei punti di forza e di debolezza, offre una prima panoramica che rivela una buona congruenza tra i termini utilizzati dalle scuole e gli specifici campi aperti entro cui si collocano tali segmenti di testo. In altre parole le scuole hanno utilizzato lemmi che rappresentano gli elementi di forza e di debolezza. In linea generale le istituzioni scolastiche manifestano una adeguata padronanza degli argomenti che, dal punto di vista semantico, risiedono nella dimensione oggetto di indagine.

A tal proposito è stata effettuata una ulteriore analisi delle partizioni basata sul confronto con i termini presenti nelle domande guida di questa sezione del RAV. L'obiettivo è quello rilevare informazioni sul livello di specificità dei termini impiegati (cfr. *La metavalutazione del RAV: obiettivi e metodologia*, pag. 11).

La procedura ha evidenziato che, tra i punti di forza, il 18% delle parole rappresenta la quota delle parole sovrautilizzate (aventi gli *Z score* >2. Cfr. Glossario); tra di esse il 5,3% è contenuta nella fascia di frequenza "alta". Si tratta di valori relativamente alti in considerazione del fatto che il numero di parole presenti nelle domande guida è molto inferiore. La medesima procedura applicata ai punti di debolezza ha fatto registrare una quota irrilevante di parole sovrautilizzate (*Z score* > 2: 0,7% dei termini presenti).

Tali evidenze confermano, come già in precedenza indicato, che le scuole hanno argomentato in maniera più ampia gli elementi positivi che hanno messo in campo per garantire ai propri studenti maggiori possibilità di raggiungimento del successo scolastico. Diversamente, nella descrizione degli elementi negativi si sono riferite per lo più agli stimoli presenti nelle domande guida, probabilmente con l'obiettivo più o meno esplicito di 'parlare meno' di tali

aspetti, come abbiamo già avuto modo di evidenziare

2.3 L'analisi testuale: i temi emergenti

L'analisi testuale sui punti di forza e di debolezza è stata condotta con lo scopo di approfondire la conoscenza del corpus, già in parte esplorato, anche attraverso l'individuazione di nuclei tematici peculiari.

Dopo aver operato una prima osservazione sulle singole parole attraverso l'analisi lessicale, si è voluto approfondire lo studio delle associazioni tra le parole con lo scopo di operare un processo di categorizzazione dei termini e ridurre la complessità dei testi ad alcune dimensioni latenti sottostanti.

In questa fase pertanto, l'obiettivo primario del lavoro consiste nell'attribuzione di senso alle parole chiave riconosciute come tali (*keyword*, Cfr. Glossario), rispetto alle altre parole a esse associate.

A tal fine è stata utilizzata la tecnica *Topic extraction* di Wordstat che, attraverso una analisi in componenti principali, ha permesso di individuare dei nuclei tematici peculiari e le relative parole chiave. Nello specifico la procedura ha considerato come unità di analisi le singole frasi.⁸

I risultati emersi hanno consentito di costruire un vocabolario tematico che esprime una sintesi dei temi presenti, nonché di avere una visione più puntuale della struttura complessiva dei documenti.

Focalizzando l'attenzione sulla la procedura di *Topic extraction* (Cfr. Glossario) applicata e, avendo inserito la *stop list*, è stato possibile analizzare il 79% delle parole della partizione dei punti di

⁸ L'analisi in componenti principali effettuata ha utilizzato una rotazione degli assi (Warimax) e un

peso fattoriale minimo pari a 0,40; inoltre si è deciso di limitare le occorrenze al minimo di 10%.

forza e il 72% delle parole della partizione dei punti di debolezza.

Per quel che riguarda i punti di forza, le componenti individuate rappresentano gli argomenti che sono stati trattati dalle scuole nel campo aperto dedicato. Sono state fatte differenti prove di estrazione dei *fattori* per arrivare a una estrazione esaustiva e significativa con la presenza di 10 fattori (Tab. 1.1); ciò ha consentito di individuare dei nuclei tematici specifici che, mediante l'individuazione di parole chiave che co-occorrono nella stessa frase, ben

esplicitano gli argomenti principali trattati dalle scuole.

I *topic* individuati sono stati esportati mediante una procedura automatizzata prevista dal software, al fine di strutturare un modello di dizionario (specifico per i punti di forza di questa dimensione), in quanto la distribuzione delle *keyword* è risultata essere ben differenziata rispetto ai temi emersi.

Tabella 2.1. - Temi estratti con la procedura *Topic extraction* dell'area Successo scolastico (Punti di forza)

TOPIC	KEYWORDS	EIGENVALUE	% VAR	FREQ	CASI	% CASI
REGIONALI E NAZIONALI; DATI PROVINCIALI	NAZIONALI; PROVINCIALI; REGIONALI; SUPERIORE; DATI; PROVINCIALE; REGIONALE; LINEA; PERCENTUALE	4,22	1,21	1365	492	67,9%
CRITERI	CRITERI; FORMATIVO; SUCCESSO; GARANTIRE; VALUTAZIONE; ADEGUATI; ADOTTATI	5,02	0,88	1840	456	62,9%
DISTRIBUZIONE; VOTO ALL'ESAME	DISTRIBUZIONE; VOTO; CONSEGUITO; FASCIA; ANALIZZANDO; ESAME; COLLOCA; FASCE; CONSISTENTE	3,66	0,93	1146	388	53,5%
6 E 7; 9 E 10	6; 9; 7; 2; 3; 8; 1; 5; 10; 4; 21; 0	6,52	1,32	1158	329	45,4%
PRIMARIA	PRIMARIA; SECONDARIA	3,73	1,19	491	289	39,9%
LICEO SCIENTIFICO; GIUDIZIO	LICEO; GIUDIZIO; INDIRIZZI; SCIENTIFICO; CLASSICO; BIENNIO; INDIRIZZO	5,24	1,23	380	198	27,3%
ATTIVITÀ DI RECUPERO	RECUPERO; ATTIVITÀ	22,41	1,51	201	164	22,6%
ITALIANO E MATEMATICA; LINGUE STRANIERE	MATEMATICA; LINGUE; STRANIERE; ITALIANO; CONCENTRANO; LINGUA; DISCIPLINE; DEBITI	3,27	0,91	253	126	17,4%
ESPERIENZE; FORMATIVE	ESPERIENZE; FORMATIVE; ADEGUATE; TALI; LICENZIATI; TRIENNIO	3,35	0,68	119	80	11,0%
PARTENZA	PARTENZA; PROGRESSI; CONTO; IMPEGNO	3,24	0,86	104	66	9,1%

Fonte: elaborazione INVALSI, dati RAV a.s. 2014/2015

L'ordinamento rispetto alla percentuale di casi offre una immediata visione dei temi maggiormente trattati dalle scuole: se esaminiamo i valori percentuali più elevati riscontriamo che il 68% circa dei documenti elaborati dalle scuole riguarda il confronto con i riferimenti territoriali esplicitato attraverso il ricorso a *keyword* come "nazionali; provinciali; regionali; dati; superiore; linea; percentuale". Medesima interpretazione si può dare per le parole chiave associate al *topic* "criteri" che sono presenti nel 63% dei testi e che contemplano tematiche legate sia ai "criteri

di valutazione" adottati dalle scuole sia al "successo formativo" degli alunni.

La componente denominata "distribuzione; voto all'esame", presente in più della metà dei frammenti di testo esaminati, si riferisce alla distribuzione degli studenti in fasce di livello rispetto ai risultati ottenuti all'Esame di Stato.

Nel complesso tutte le componenti individuate sono state oggetto di una rielaborazione che ha consentito di costruire il vocabolario tematico che verrà presentato nel paragrafo successivo.

2.4 Il vocabolario tematico dei punti di forza

La costruzione del vocabolario tematico ha preso l'avvio dall'analisi in componenti principali sopra descritta e si è sviluppata attraverso le seguenti ulteriori operazioni:

1. il consolidamento dei temi semantici maggiormente significativi attraverso l'associazione di parole chiave e termini caratterizzanti;

2. la rimozione di parole meno rilevanti perché comuni a più temi;
3. l'aggiunta di nuove parole ritenute caratterizzanti a seguito dell'analisi dei segmenti ripetuti e delle co-occorrenze.⁹

Il vocabolario tematico elaborato presenta un ottimo livello di copertura ed è articolato in 8 categorie (Tabella 1.2); dopo aver applicato la *stop list*, esso copre il 75% delle frasi e il 99,7 % dei documenti.

Tabella 2.2 - Frequenza e diffusione delle categorie del vocabolario tematico dei punti di forza (Successo scolastico)

CATEGORIE	FREQ.	% FREQ.	N. CASI	% CASI	TF • IDF
CONFRONTI TERRITORIALI	4616	38,5%	646	89,1%	231,3
CONFRONTO INTERNO SCUOLA	2275	19,0%	565	77,9%	246,4
CRITERI DI VALUTAZIONE	1539	12,9%	459	63,3%	305,5
FASCIA DI VOTO	1518	12,7%	422	58,2%	356,8
TRASFERIMENTI E ABBANDONI	956	8,0%	387	53,4%	260,6
DEBITI FORMATIVI	494	4,1%	221	30,5%	254,9
ATTIVITÀ DI RECUPERO	388	3,2%	217	29,9%	203,3
VOTI DEGLI STUDENTI	190	1,6%	158	21,8%	125,7

Fonte: elaborazione INVALSI, dati RAV a.s. 2014/2015

La categoria **confronti territoriali** si riferisce appunto al confronto dell'andamento degli esiti degli scrutini e delle percentuali di trasferimenti e abbandoni degli studenti della propria scuola, in relazione con le altre scuole collocate sul territorio. Ricorrono pertanto parole chiave come nazionale, provinciale, regionale, ma anche termini di confronto quali: media, in linea, superiore, inferiore.

Le *keyword* associate a questa categoria ricorrono con una frequenza

elevata e sono diffuse nell'89,1% dei documenti analizzati. In tale categoria sono stati associati due temi individuati dalla *topic*, "nazionali, provinciali" e "voti all'esame" e sono state aggiunte ulteriori parole aventi il medesimo carattere semantico.

Nel 78% dei documenti redatti dalle scuole sono presenti termini che possono essere ricondotti alla categoria **confronto**

⁹ L'analisi delle co-occorrenze consente di studiare le associazioni tra parole, individuando quelle parole che compaiono più spesso vicine tra loro.

interno scuola e che sono inerenti le classi e gli indirizzi su cui le scuole articolano i punti di forza per evidenziare situazioni per loro positive. È caratterizzata dalla presenza di parole chiave che esprimono differenze tra gli indirizzi di scuola (classico, scientifico), e ulteriori parole tra le classi all'interno di un medesimo ordine di scuola.

Nella categoria **fascia di voto** si trovano i termini relativi alla collocazione degli studenti e delle classi rispetto ai voti o agli esami. Tale categoria contiene la maggior parte delle *keyword* associate al *topic* "distribuzione" e risultano diffuse nel 60% dei documenti compilati. Le scuole evidenziano elementi che le caratterizzano rispetto ad altre situazioni:

Analizzando la distribuzione degli studenti per fascia di voto conseguito all'Esame di Stato, si rilevano alcune eccellenze, che confermano il trend positivo, per quanto riguarda la motivazione allo studio degli studenti e la crescente attenzione della scuola nel garantire il successo formativo [caso 523].

La categoria **criteri di valutazione** rappresenta uno dei temi individuati e presente in oltre la metà dei documenti analizzati (63,3%). Trattandosi di punti di forza, le *keyword* a essa associate mostrano l'adeguatezza dei criteri adottati dalle scuole come contributo per il raggiungimento del successo formativo degli studenti. Si riporta a titolo esemplificativo il seguente caso:

I criteri di valutazione adottati dalla scuola sono adeguati a garantire il successo formativo degli studenti, in quanto mettono sempre al centro del processo di insegnamento-apprendimento lo studente valutandone la situazione di partenza ed il percorso fatto durante l'anno scolastico [caso 186].

Nella categoria **voti degli studenti** si concentrano i caratteri numerici che rappresentano prioritariamente i voti ottenuti dagli studenti agli esami; tuttavia dall'analisi delle concordanze è emerso che in alcuni casi i numeri si riferiscono ai punteggi percentuali riportati dalle scuole per indicare la distribuzione di studenti rispetto alle votazioni ottenute. L'esempio successivo ben rappresenta tali casi:

Dall'analisi delle votazioni conseguite all'Esame di Stato, risulta che la fascia medio-bassa (fascia 6-7) è, in percentuale, leggermente più ampia della fascia medio-alta; la fascia del 10 è del 14,6% ed è in percentuale più ampia rispetto alla media provinciale, regionale e nazionale [caso 431].

Le categorie denominate **trasferimenti e abbandoni** e **debiti formativi** sono presenti in misura minore nei documenti: rispettivamente nel 53% e nel 30% di essi. In questi casi le parole associate a queste categorie sono espresse in termini positivi, pertanto si collocano sempre accanto a parole come *numero esiguo, bassa percentuale, pochi casi*.

Infine la categoria **attività di recupero** comprende termini meno frequenti (circa il 30% dei documenti), relativi a varie forme di interventi a sostegno degli studenti in difficoltà.

Nel complesso, la maggior parte delle scuole hanno redatto dei testi sintetici che riprendono le frasi presenti nei RAV, come i temi presenti nella definizione delle aree e nelle domande guida; in alcuni casi esse realizzano un lavoro di rielaborazione più complesso e argomentano i temi correlandoli sia alla normativa di riferimento sia al processo complessivo di pratiche educative e didattiche da loro attuato. La tabella successiva (1.3) offre la distribuzione complessiva delle *keyword* associate a tutte le categorie sopra descritte.

Tabella 2.3 - Frequenza e diffusione delle *keyword* nelle categorie del vocabolario tematico dei punti di forza (Successo scolastico)

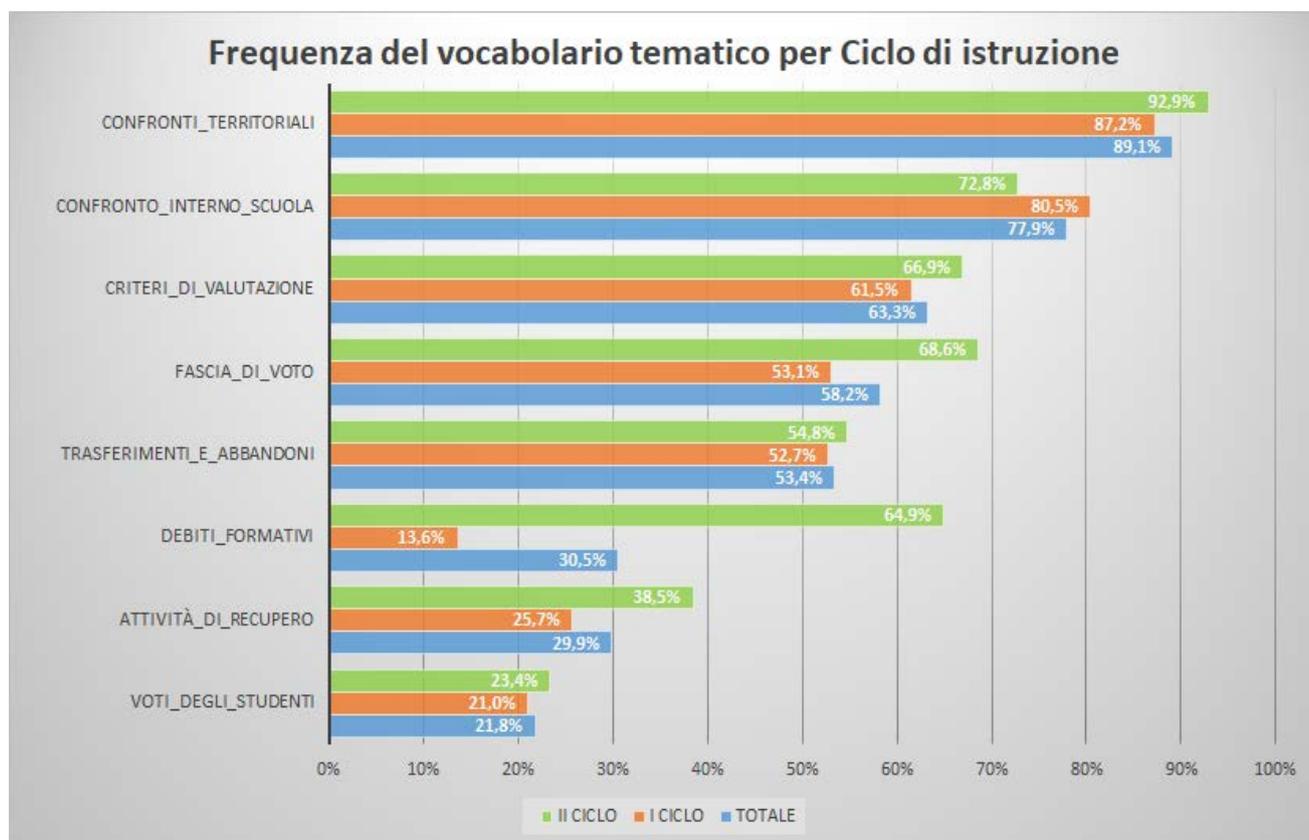
CONFRONTI TERRITORIALI	FREQUENCY	% CASES	CRITERI DI VALUTAZIONE	FREQUENCY	% CASES	FASCIA DI VOTO	FREQUENCY	% CASES
AMMESSI	593	59,9%	CRITERI_DI_VALUTAZIONE	278	37,9%	ESAME	297	36,4%
PERCENTUALE	585	45,9%	FORMATIVO	215	27,6%	FASCIA	186	18,6%
MEDIA	544	41,1%	VALUTAZIONE	203	21,7%	VOTO	178	16,4%
CLASSE_SUCESSIVA	483	54,5%	ADOTTATI	173	23,7%	FASCE	161	18,2%
RISPETTO	358	30,3%	ADEGUATI	172	23,6%	DISTRIBUZIONE	154	18,6%
CLASSE	262	22,9%	GARANTIRE_IL_SUCESSO	171	23,5%	LIVELLO	104	12,6%
DATI	261	24,3%	SUCCESSO_FORMATIVO	134	17,2%	CONSEGUITO	99	13,1%
SUPERIORE	233	23,3%	SUCCESSO	62	7,3%	MEDIO	99	11,9%
DATO	228	18,8%	CRITERI	52	6,2%	FASCIA_DI_VOTO	93	11,9%
REGIONALE	198	19,6%	GARANTIRE	37	4,7%	COLLOCA	72	8,7%
LINEA	196	21,9%	CONDIVISI	31	4,3%	0	61	5,9%
NAZIONALI	181	18,3%	CONDIVISIONE	11	1,5%	ANALIZZANDO	41	5,7%
PROVINCIALE	147	13,8%	VOTI DEGLI STUDENTI	FREQUENCY	% CASES	CONSISTENTE	34	4,7%
INFERIORE	127	13,4%	10	155	15,9%	ATTIVITÀ DI RECUPERO	FREQUENCY	% CASES
REGIONALI	126	12,8%	7	152	14,6%	RECUPERO	106	13,0%
2013	99	10,5%	8	151	15,3%	ATTIVITÀ	94	10,8%
PROVINCIALI	87	8,3%	6	142	14,1%	INTERVENTI	61	7,3%
14	57	6,2%	9	139	15,0%	CORSI	40	5,1%
CONFRONTO INTERNO SCUOLA	FREQUENCY	% CASES	LODE	83	11,2%	CORSI_DI_RECUPERO	37	4,7%
SCUOLA	1074	68,1%	90	48	6,1%	STRATEGIE	31	4,1%
SECONDARIA	377	33,4%	SCRUTINI	45	5,9%	PROVE_DI_VERIFICA	12	1,4%
PRIMARIA	365	34,3%	80	44	5,7%	DEBITI FORMATIVI	FREQUENCY	% CASES
GRADO	239	25,1%	CONSEGUITI	38	4,8%	GIUDIZIO	116	13,4%
LICEO	112	8,0%	91	32	4,4%	BIENNIO	101	8,8%
SCIENTIFICO	43	3,9%	81	23	3,0%	DISCIPLINE	77	8,7%
ORDINI_DI_SCUOLA	38	4,7%	ESITO	18	2,3%	MATEMATICA	53	6,8%
CLASSICO	27	2,2%	TRASFERIMENTI E ABBANDONI	FREQUENCY	% CASES	DEBITI_FORMATIVI	32	4,3%
			TRASFERIMENTI	229	26,1%	FORMATIVI	26	3,3%
			ABBANDONI	214	27,2%	SOSPENSIONE_DEL_GIUDIZIO	25	3,2%
			USCITA	194	24,6%	SOSPENSIONE	23	3,2%
			ENTRATA	152	19,5%	CONCENTRANO	23	3,2%
			CORSO_D'ANNO	100	12,0%	DEBITI	18	2,5%
			ALTRE_SCUOLE	38	5,2%			
			STUDI_IN_CORSO	27	3,7%			

Fonte: elaborazione INVALSI, dati RAV a.s. 2014/2015

Il vocabolario tematico fin qui descritto è stato esaminato anche rispetto alla variabile relativa al ciclo di istruzione. L'osservazione dei risultati, presentati nel grafico successivo, ha fatto rilevare una differenza statisticamente significativa per tre delle categorie sopra descritte: **debiti formativi, attività di recupero e fascia di voto**¹⁰.

Si evidenzia una tendenza delle scuole del II ciclo a trattare in misura prevalente argomenti connessi a tali categorie. Ciò è da imputare probabilmente anche al fatto che esse sono tenute a implementare corsi di recupero e attività di sostegno che vanno a costituire una parte ordinaria e permanente del piano dell'offerta formativa.¹¹

Figura 2.5 – Distribuzione delle categorie del vocabolario tematico (punti di forza) per Ciclo di istruzione



Fonte: elaborazione INVALSI, dati RAV a.s. 2014/2015

¹⁰ I confronti sono stati effettuati con l'applicazione del test del χ^2 ("chi quadro").

¹¹ Secondo quanto definito nel Decreto Ministeriale 80 del 3 ottobre 2007, ogni scuola deve impegnarsi nel corso dell'anno a offrire ai propri alunni delle proposte didattiche ed educative volte "a far superare agli studenti le insufficienze che rischiano di compromettere il proseguimento dei loro studi".

2.5 L'analisi testuale dei punti di debolezza

Le medesime procedure impiegate per l'analisi dei punti di forza sono state replicate sui segmenti di testo relativi ai punti di debolezza. Anche in questa circostanza, il punto di partenza del lavoro è costituito dalla *Topic extraction*, implementata al fine di individuare alcuni temi principali attorno ai quali si associano le parole chiave aventi il medesimo potenziale semantico. La procedura ha preso in considerazione il 72% dei testi presenti. I risultati rivelano diverse analogie con i temi emersi nei punti di forza della medesima dimensione.

Se osserviamo le percentuali di occorrenze (Tabella 1.5), il tema ricorrente in più della metà dei testi presenti (56%) riguarda i valori numerici che le scuole riportano per effettuare confronti e che rappresentano per lo più i "voti degli studenti"; a seguire troviamo il *topic*

relativo alla collocazione delle scuole rispetto alle diverse "fasce di voto" presenti (42%).

Seguono poi le componenti relative a materie e indirizzi (32%): si tratta di un *topic* particolarmente rilevante in quanto a esso si associano *keyword* che rappresentano gli elementi di criticità che le scuole denunciano e dove attivano strategie di intervento per migliorare il successo scolastico degli studenti.

Il *topic* denominato "criteri, successo formativo" è presente nel 30% dei documenti esaminati ed è quello che contiene *keyword* riferite ai temi principali dell'area del RAV; appare interessante ricordare che il medesimo *topic* estratto per i punti di forza era presente in più del doppio dei documenti (62%).

Altri documenti (circa il 18%) contengono termini legati alle possibili cause e motivi alla base delle disuguaglianze e difficoltà che gli studenti incontrano nel proprio percorso formativo.

Tabella 2.5 - Temi estratti con la procedura *Topic extraction* dell'area Successo scolastico (Punti di debolezza)

TOPIC	KEYWORDS	EIGENVALUE	% VAR	FREQ	CASI	% CASI
13; 2,5	3; 5; 2; 4; 8; 1; 6; 7; CLASSE; MEDIA; 10	6,39	1,65	1133	407	56,1%
VOTO; ESAME	VOTO; ESAME; DISTRIBUZIONE; CONSEGUITO; FASCE; FASCIA; COLLOCA; ANALIZZANDO	3,72	1,14	771	304	41,9%
MATEMATICA; DEBITI FORMATIVI	MATEMATICA; DEBITI; FORMATIVI; DISCIPLINE; CONCENTRANO; INGLESE; MATERIE; INDIRIZZO; BIENNIO; FISICA; ITALIANO	5,83	1,52	715	235	32,4%
CRITERI; SUCCESSO FORMATIVO	CRITERI; SUCCESSO; FORMATIVO; VALUTAZIONE; ADOTTATI; ADEGUATI; GARANTIRE	3,31	0,96	454	222	30,6%
CAUSE; METODO	CAUSE; METODO; SCOLASTICA; SCARSO; SOCIO; PREPARAZIONE	3,52	1,13	169	128	17,7%
REGIONALI	REGIONALI; NAZIONALI; PROVINCIALI; MEDIE	3,17	0,89	231	123	17,0%
ATTIVITÀ DI RECUPERO	RECUPERO; ATTIVITÀ; RISORSE	20,33	1,55	171	118	16,3%
61-70; LICEO	70; 61; LICEO; 100; 80; TECNICO; 91	4,04	1,39	182	96	13,2%
DOVUTI	DOVUTI; STRANIERI; SPOSTAMENTI	3,64	1,29	106	76	10,5%
2013/2014	2014; 2013; 2015	3,09	1,02	85	46	6,3%

Fonte: elaborazione INVALSI, dati RAV a.s. 2014/2015

2.6 Il vocabolario tematico dei punti di debolezza

L'elaborazione del vocabolario tematico dei punti di debolezza ha ripercorso le tappe già individuate per l'analisi dei punti di forza al fine di consolidare i temi semantici più significativi attraverso:

- ✓ l'associazione di parole caratterizzanti;
- ✓ la rimozione di parole ambigue o comuni a più temi;
- ✓ l'aggiunta di termini caratterizzanti il *topic* individuato.

Il vocabolario tematico elaborato presenta un buon livello di copertura ed è articolato in 7 categorie distinte (Tabella 1.6); dopo aver applicato la *stop list*, esso copre il 60% delle frasi e il 91% dei documenti. Le categorie sono

pressoché equivalenti a quelle individuate nei punti di forza; tale tendenza era prevedibile in considerazione del fatto che gli argomenti trattati dalle scuole sono i medesimi, seppur espressi in termini positivi o negativi. Nonostante ciò, per alcune categorie, le parole a esse associate risultano avere numerosità inferiori.

Differentemente dal vocabolario dei punti di forza, in questo relativo a quelli di debolezza non è presente la categoria **confronto interno scuola**; infatti, le parole a essa associate si collocano in diverse categorie senza generare un nucleo semantico indipendente.

Ciò potrebbe risiedere nel fatto che nella diagnosi dei propri punti di debolezza le scuole evidenziano carenze o problematiche presenti trascurando il confronto tra cicli scolastici e indirizzi di studio.

Tabella 2.6 - Frequenza e diffusione delle categorie del vocabolario tematico dei punti di debolezza (Successo scolastico)

CATEGORIE	FREQ.	% FREQ.	N. CASI	% CASI	TF • IDF
CONFRONTI TERRITORIALI	2189	35,6%	502	69,2%	349,4
VOTI DEGLI STUDENTI	1042	17,9%	266	36,7%	453,7
DEBITI FORMATIVI	995	17,0%	284	39,2%	405
FASCIA DI VOTO	966	16,6%	330	45,5%	330,2
CRITERI DI VALUTAZIONE	414	7,1%	222	30,6%	212,8
TRASFERIMENTI E ABBANDONI	349	6,0%	203	28,0%	192,9
ATTIVITÀ DI RECUPERO	196	3,4%	118	16,3%	154,5

Fonte: elaborazione INVALSI, dati RAV a.s. 2014/2015

Le *keyword* presenti in ciascuna categoria risultano assimilabili a quelle già esposte nel paragrafo precedente. È interessante notare però la connotazione differente con cui vengono presentate.

Nella categoria **criteri di valutazione** ad esempio, si collocano frasi che manifestano la difficoltà delle scuole nel condividere dei criteri di valutazione uniformi tra le classi e i plessi presenti; a parere delle scuole, tale difformità diviene un elemento di criticità che incide negativamente sul successo scolastico degli alunni. Gli esempi riportati di seguito aiutano a comprendere meglio tali contesti:

Le valutazioni degli alunni presentano caratteri di disomogeneità tra classi e sezioni. Benché i criteri di valutazione siano comuni, la mancanza di più frequenti riferimenti a prove oggettive e comuni tra le classi, rende l'attribuzione dei voti – soprattutto in sede di scrutinio finale – soggetta a notevole variabilità e discrezionalità da parte dei docenti [caso 552].

Permane una certa difficoltà nella condivisione dei criteri di valutazione tra ordini di scuola diversi e tra plessi differenti [caso 70].

In generale la descrizione delle *keyword* associate alle categorie dei punti di forza si può sovrapporre in maniera quasi simmetrica alle categorie dei punti di debolezza in quanto non si evidenziano differenze rilevanti tra i nuclei semantici individuati.

Tabella 2.6 - Frequenza e diffusione delle *keyword* nelle categorie del vocabolario tematico dei punti di debolezza (Successo scolastico)

CONFRONTI TERRITORIALI	FREQUENCY	% CASES	DEBITI FORMATIVI	FREQUENCY	% CASES	FASCIA DI VOTO	FREQUENCY	% CASES
PERCENTUALE	356	32,0%	MATEMATICA	132	14,9%	ESAME	199	23,9%
CLASSI	313	17,4%	BIENNIO	125	11,5%	FASCIA	162	15,3%
CLASSE	303	28,1%	DISCIPLINE	122	14,3%	FASCE	119	12,8%
MEDIA	301	26,2%	DEBITI	92	10,9%	VOTO	119	12,8%
RISPETTO	263	24,6%	INDIRIZZO	88	8,7%	DISTRIBUZIONE	91	12,1%
SUPERIORE	126	13,8%	FORMATIVI	81	10,3%	LIVELLO	72	8,6%
INFERIORE	90	10,8%	CONCENTRANO	73	9,7%	CONSEGUITO	68	9,1%
LICEO	89	6,3%	INGLESE	45	5,5%	FASCIA_DI_VOTO	57	7,2%
NAZIONALI	85	9,9%	MATERIE	45	5,5%	COLLOCA	52	6,5%
TECNICO	76	6,1%	SCOLASTICA	42	5,1%	ANALIZZANDO	27	3,6%
MEDIE	74	8,4%	ITALIANO	38	4,4%			
REGIONALI	56	7,3%	PREPARAZIONE	24	3,0%			
PROVINCIALI	38	4,7%	FISICA	23	2,9%			
MEDIE DI RIFERIMENTO	19	2,1%	SCARSO	14	1,7%			
VOTI DEGLI STUDENTI	FREQUENCY	% CASES	METODO_DI_STUDIO	14	1,7%			
6	160	17,1%	CAUSE	12	1,7%	ATTIVITÀ DI RECUPERO	FREQUENCY	% CASES
7	133	13,4%	DISPERSIONE_SCOLASTICA	11	1,5%	RECUPERO	69	8,0%
3	112	10,2%	CRITERI DI VALUTAZIONE	FREQUENCY	% CASES	ATTIVITÀ	59	7,6%
10	105	10,8%	VALUTAZIONE	148	15,3%	RISORSE	48	5,8%
5	95	8,3%	CRITERI_DI_VALUTAZIONE	65	8,1%	CORSI_DI_RECUPERO	20	2,3%
8	94	10,6%	SUCCESSO_FORMATIVO	47	6,2%			
4	89	7,3%	GARANTIRE	40	5,5%			
9	78	8,6%	FORMATIVO	27	3,6%	TRASFERIMENTI E ABBANDONI	FREQUENCY	% CASES
LODE	61	8,1%	ADOTTATI	26	3,6%	TRASFERIMENTI	149	16,7%
100	39	3,9%	CRITERI	24	3,3%	ABBANDONO	72	9,4%
70	27	3,0%	ADEGUATI	22	2,9%	STRANIERI	68	7,2%
61	18	2,1%	SUCCESSO	15	1,9%	DOVUTI	42	5,1%
91	17	2,2%			SPOSTAMENTI	18	2,3%	

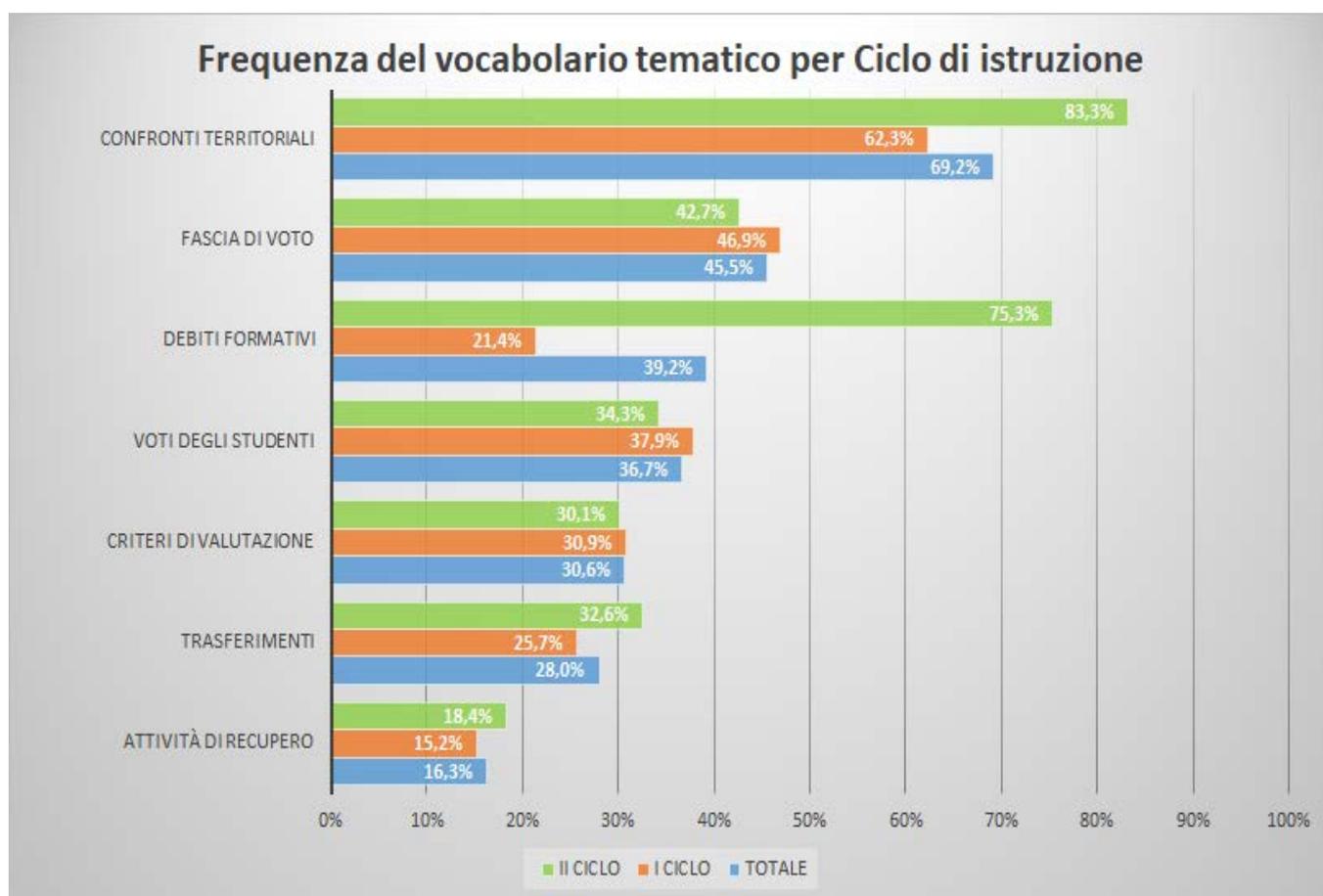
Fonte: elaborazione INVALSI, dati RAV a.s. 2014/2015

Se si osserva la distribuzione delle categorie del vocabolario tematico rispetto ai due cicli di istruzione (Grafico 1.2), si rilevano differenze statisticamente significative nelle categorie relative ai **debiti formativi** e ai **confronti territoriali**¹².

Nel primo caso lo scarto rilevato è consistente: le scuole del II ciclo trattano in

misura nettamente superiore i temi legati alle difficoltà riscontrate dagli studenti in alcune materie (ad esempio in Matematica o più in generale in alcune discipline, principalmente nel primo biennio). Diversamente le scuole del I ciclo descrivono nei punti di debolezza aspetti legati al confronto della propria situazione rispetto ai riferimenti territoriali disponibili nel RAV.

Figura 2.6 - Distribuzione delle categorie del vocabolario tematico (punti di debolezza) per Ciclo di istruzione



Fonte: elaborazione INVALSI, dati RAV a.s. 2014/2015

¹² I confronti sono stati effettuati con l'applicazione del test del χ^2 ("chi quadro").

RISULTATI IN SINTESI

- L'analisi testuale condotta sulla dimensione del Successo scolastico del RAV ha offerto diversi elementi utili a rispondere alle domande e alle ipotesi di ricerca che hanno orientato il percorso di analisi. Nella redazione dei campi aperti dei punti di forza e dei punti di debolezza le scuole hanno avviato un lavoro di riflessione rispetto ai temi presentati negli indicatori disponibili per questa area. Si tratta di indicatori che vengono restituiti alle scuole a livello centrale e che contengono dati oggettivi che rappresentano le diverse situazioni rispetto agli elementi che contribuiscono, o meno, a realizzare il successo scolastico degli studenti.
- In questa dimensione del RAV le scuole hanno fatto un esplicito riferimento alla terminologia correntemente in uso e ormai consolidata, proveniente dalle domande guida e dai contenuti presenti nei documenti, discostandosi in maniera lieve dai temi e dalle parole che hanno stimolato la riflessione.
- Dal punto di vista metodologico, l'analisi lessicale e l'analisi testuale, hanno permesso di analizzare i testi in maniera sistematica: dal semplice conteggio delle occorrenze è stato possibile individuare il grado di diffusione di specifici nuclei semantici e tematici nel corpus testuale esaminato.
- Le diverse tecniche statistiche impiegate nelle varie fasi del lavoro, hanno fornito risultati coerenti tra di loro e le informazioni estratte in ciascuna fase sono state confermate dalle analisi successivamente effettuate: analisi lessicale, ACP, analisi dei segmenti ripetuti e delle co-occorrenze (a supporto del consolidamento dei nuclei tematici).
- In sintesi, le tendenze emerse dalle analisi sono sostanzialmente le seguenti:
 - un maggior impegno dedicato dalle scuole nell'articolare i testi relativi ai punti di forza rispetto ai punti di debolezza, sia in termini quantitativi (maggiore numero di parole utilizzate); sia in termini qualitativi (ricorso a parole originali in misura maggiore rispetto a quelle presenti nelle domande guida);
 - differente utilizzo delle categorie del vocabolario rispetto al ciclo di appartenenza: nel complesso le scuole del I ciclo hanno fatto rilevare uno scarto positivo statisticamente significativo in termini di varietà di *keyword* utilizzate e ascrivibili alle specifiche categorie.

RIFLESSIONI E PROSPETTIVE

- Il RAV ha favorito un percorso di autovalutazione all'interno di un *frame* definito, sarebbe auspicabile approfondire ulteriormente la coerenza tra i punti di forza e di debolezza individuati e le motivazioni legate al giudizio autoattribuito.
- La trasparenza delle scelte metodologiche e delle procedure che hanno guidato le analisi, consentono di poter replicare le stesse su altri segmenti di testo di altre aree del RAV permettendo una metavalutazione completa dello strumento.
- L'analisi della coerenza tra punti di forza e di debolezza con le priorità e gli obiettivi individuati dalle scuole si delinea come tassello conclusivo di un percorso di ricerca che ha approfondito tutti gli aspetti provenienti dalle ipotesi di ricerca definite in origine.

Bibliografia

Bolasco, S., (1999). *Analisi multidimensionale dei dati. Metodi, strategie e criteri d'interpretazione*. Roma, Carocci.

Bolasco, S., (2013). *L'analisi automatica dei testi. Fare ricerca con il text mining*. Roma, Carocci.

Giuliano, L., & La Rocca, G. (2008). *L'analisi automatica e semiautomatica dei dati testuali*. Milano: LED Edizioni universitarie di Lettere Economia Diritto.

Lebart L., Salem A., Berry L., (1998) *Exploring Textual Data*. Kluwer Academic Publishers, Dordrecht-Boston-London.

Appendice delle tabelle

Tabella 2.7 - Parole a frequenza elevata presenti nei punti di forza dell'area Successo scolastico

	FREQ.	% FREQ.	N. CASI	% CASI	TF • IDF
DI	4182	5,3%	700	96,6%	63,7
E	2190	2,8%	623	85,9%	144,2
IN	1690	2,2%	557	76,8%	193,5
LA	1549	2,0%	586	80,8%	143,2
IL	1420	1,8%	541	74,6%	180,5
SCUOLA	1111	1,4%	503	69,4%	176,4
PER	1071	1,4%	474	65,4%	197,7
DEGLI	1056	1,3%	512	70,6%	159,5
ALLA	1052	1,3%	525	72,4%	147,5
STUDENTI	1046	1,3%	471	65,0%	195,9
I	1025	1,3%	521	71,9%	147,1
SONO	927	1,2%	468	64,6%	176,2
NON	899	1,1%	457	63,0%	180,2
È	875	1,1%	432	59,6%	196,7
CHE	856	1,1%	408	56,3%	213,7
SI	833	1,1%	446	61,5%	175,8
CON	830	1,1%	415	57,2%	201,1
ALUNNI	800	1,0%	398	54,9%	208,4
A	760	1,0%	437	60,3%	167,1
CLASSE	746	1,0%	459	63,3%	148,1
GLI	699	0,9%	415	57,2%	169,4
DEL	671	0,9%	355	49,0%	208,1
LE	637	0,8%	340	46,9%	209,5
NELLA	621	0,8%	322	44,4%	218,9
AMMESSI	592	0,8%	434	59,9%	131,9
PERCENTUALE	585	0,7%	333	45,9%	197,7
DELLA	583	0,7%	336	46,3%	194,7
AL	578	0,7%	335	46,2%	193,8
MEDIA	544	0,7%	298	41,1%	210,1
L	543	0,7%	326	45,0%	188,5
UN	521	0,7%	318	43,9%	186,5
UNA	509	0,7%	336	46,3%	170
ANNO	496	0,6%	316	43,6%	178,9
SUCCESSIVA	487	0,6%	399	55,0%	126,3
VALUTAZIONE	480	0,6%	362	49,9%	144,8
NAZIONALE	452	0,6%	267	36,8%	196,1
NEL	452	0,6%	279	38,5%	187,5
DEI	446	0,6%	294	40,6%	174,8
ALL	419	0,5%	318	43,9%	150
CLASSI	404	0,5%	230	31,7%	201,4
DELL	384	0,5%	262	36,1%	169,7
SECONDARIA	377	0,5%	242	33,4%	179,6
DELLE	372	0,5%	247	34,1%	174
SUCCESSO	367	0,5%	321	44,3%	129,9
PRIMARIA	364	0,5%	249	34,3%	168,9

RISPETTO	358	0,5%	220	30,3%	185,4
FORMATIVO	349	0,4%	305	42,1%	131,2
STATO	337	0,4%	296	40,8%	131,1
CRITERI	330	0,4%	308	42,5%	122,7
PIÙ	324	0,4%	217	29,9%	169,7
SCOLASTICO	304	0,4%	234	32,3%	149,3
ESAME	297	0,4%	264	36,4%	130,3
ISTITUTO	291	0,4%	205	28,3%	159,6
DA	285	0,4%	211	29,1%	152,8
NELLE	281	0,4%	183	25,2%	168
FASCIA	279	0,4%	190	26,2%	162,3
CORSO	277	0,4%	211	29,1%	148,5
ALLE	275	0,4%	196	27,0%	156,2
VOTO	270	0,3%	191	26,3%	156,4
DATI	260	0,3%	176	24,3%	159,9
NUMERO	258	0,3%	178	24,6%	157,4
DALLA	244	0,3%	216	29,8%	128,3
GRADO	237	0,3%	182	25,1%	142,3
SUPERIORE	232	0,3%	168	23,2%	147,3
TRASFERIMENTI	231	0,3%	189	26,1%	134,9

Fonte: elaborazione INVALSI, dati RAV a.s. 2014/2015

Tabella 2.8 - Parole a frequenza media presenti nei punti di forza dell'area Successo scolastico

	FREQ.	% FREQ.	N. CASI	% CASI	TF • IDF
DATO	228	0,3%	136	18,8%	165,7
PRIMO	228	0,3%	181	25,0%	137,4
ANCHE	214	0,3%	163	22,5%	138,7
ABBANDONI	213	0,3%	196	27,0%	121
GARANTIRE	208	0,3%	200	27,6%	116,3
AI	199	0,3%	143	19,7%	140,3
LINEA	196	0,3%	159	21,9%	129,2
REGIONALE	195	0,3%	139	19,2%	139,9
USCITA	193	0,3%	177	24,4%	118,2
O	191	0,2%	153	21,1%	129
D	181	0,2%	151	20,8%	123,3
NAZIONALI	180	0,2%	133	18,3%	132,6
TUTTI	175	0,2%	145	20,0%	122,3
SIA	174	0,2%	114	15,7%	139,8
ADOTTATI	173	0,2%	172	23,7%	108,1
ADEGUATI	172	0,2%	171	23,6%	107,9
ED	172	0,2%	130	17,9%	128,4
RISULTATI	170	0,2%	145	20,0%	118,8
TRA	165	0,2%	125	17,2%	126
AGLI	164	0,2%	139	19,2%	117,6
CASI	164	0,2%	138	19,0%	118,2
PERCENTUALI	163	0,2%	121	16,7%	126,7
RISULTA	163	0,2%	131	18,1%	121,1
HA	162	0,2%	127	17,5%	122,6

FASCE	161	0,2%	132	18,2%	119,1
NELL	160	0,2%	129	17,8%	120
AD	154	0,2%	132	18,2%	113,9
DISTRIBUZIONE	154	0,2%	135	18,6%	112,4
ABBANDONO	153	0,2%	145	20,0%	106,9
ALTA	153	0,2%	128	17,7%	115,2
ENTRATA	152	0,2%	141	19,5%	108,1
RECUPERO	150	0,2%	123	17,0%	115,6
PROVINCIALE	145	0,2%	98	13,5%	126
RIFERIMENTO	145	0,2%	92	12,7%	130
MEDIE	144	0,2%	94	13,0%	127,8
GIUDIZIO	141	0,2%	117	16,1%	111,7
ESITI	139	0,2%	118	16,3%	109,6
HANNO	139	0,2%	108	14,9%	114,9
PARTE	130	0,2%	111	15,3%	106
INFERIORE	126	0,2%	97	13,4%	110,1
REGIONALI	126	0,2%	93	12,8%	112,4
DAL	123	0,2%	110	15,2%	100,7
ANNI	118	0,2%	97	13,4%	103,1
AMMISSIONE	114	0,1%	97	13,4%	99,6
LICEO	112	0,1%	58	8,0%	122,9
QUASI	111	0,1%	98	13,5%	96,5
QUANTO	110	0,1%	94	13,0%	97,6
DOCENTI	109	0,1%	94	13,0%	96,7
STUDI	107	0,1%	94	13,0%	94,9
LIVELLO	104	0,1%	91	12,6%	93,7
RISULTANO	104	0,1%	95	13,1%	91,8
STATI	103	0,1%	76	10,5%	100,9
PERCORSO	102	0,1%	88	12,1%	93,4
SOLO	100	0,1%	90	12,4%	90,6
SUPERIORI	100	0,1%	82	11,3%	94,7
BIENNIO	99	0,1%	64	8,8%	104,4
CONSEGUITO	99	0,1%	95	13,1%	87,4
MEDIO	99	0,1%	86	11,9%	91,7
MENTRE	97	0,1%	79	10,9%	93,4
PRIMA	97	0,1%	85	11,7%	90,3
ATTIVITÀ	94	0,1%	78	10,8%	91
BASSA	91	0,1%	76	10,5%	89,1
COMPETENZE	91	0,1%	81	11,2%	86,6
COME	90	0,1%	75	10,3%	88,7
QUELLA	90	0,1%	76	10,5%	88,2
SCOLASTICA	88	0,1%	82	11,3%	83,3
MOLTO	86	0,1%	76	10,5%	84,2
PRIME	86	0,1%	73	10,1%	85,7
PROVINCIALI	86	0,1%	60	8,3%	93,1
FAMIGLIE	85	0,1%	73	10,1%	84,7
RIGUARDA	85	0,1%	70	9,7%	86,3

<i>SCUOLE</i>	85	0,1%	74	10,2%	84,2
<i>ESAMI</i>	83	0,1%	79	10,9%	79,9
<i>LODE</i>	83	0,1%	81	11,2%	79
<i>MA</i>	83	0,1%	73	10,1%	82,8
<i>NEI</i>	83	0,1%	75	10,3%	81,8
<i>SECONDA</i>	83	0,1%	62	8,6%	88,6
<i>VALUTAZIONI</i>	82	0,1%	67	9,2%	84,8
<i>DALL</i>	80	0,1%	73	10,1%	79,8
<i>DIFFICOLTÀ</i>	80	0,1%	72	9,9%	80,2
<i>DIPLOMATI</i>	80	0,1%	66	9,1%	83,3
<i>VOTAZIONE</i>	80	0,1%	66	9,1%	83,3

Fonte: elaborazione INVALSI, dati RAV a.s. 2014/2015

Tabella 2.9 - Parole a frequenza elevata presenti nei punti di debolezza dell'area Successo scolastico

	FREQ.	% FREQ.	N. CASI	% CASI	TF • IDF
<i>DI</i>	3024	5,4%	677	93,4%	90
<i>E</i>	1478	2,6%	533	73,5%	197,5
<i>IN</i>	1260	2,2%	490	67,6%	214,4
<i>LA</i>	958	1,7%	440	60,7%	207,8
<i>IL</i>	777	1,4%	362	49,9%	234,4
<i>SI</i>	773	1,4%	434	59,9%	172,3
<i>PER</i>	712	1,3%	367	50,6%	210,5
<i>NON</i>	689	1,2%	396	54,6%	181
<i>CHE</i>	661	1,2%	360	49,7%	201
<i>A</i>	573	1,0%	328	45,2%	197,4
<i>I</i>	570	1,0%	346	47,7%	183,1
<i>STUDENTI</i>	567	1,0%	320	44,1%	201,4
<i>SCUOLA</i>	553	1,0%	332	45,8%	187,6
<i>ALUNNI</i>	540	1,0%	326	45,0%	187,4
<i>ALLA</i>	525	0,9%	299	41,2%	202
<i>È</i>	523	0,9%	291	40,1%	207,3
<i>DEL</i>	518	0,9%	306	42,2%	194,1
<i>CON</i>	506	0,9%	310	42,8%	186,7
<i>LE</i>	504	0,9%	297	41,0%	195,3
<i>DEGLI</i>	499	0,9%	320	44,1%	177,2
<i>UNA</i>	441	0,8%	295	40,7%	172,2
<i>UN</i>	435	0,8%	257	35,5%	195,9
<i>DELLA</i>	426	0,8%	270	37,2%	182,7
<i>SONO</i>	425	0,8%	277	38,2%	177,6
<i>PIÙ</i>	377	0,7%	237	32,7%	183,1
<i>NELLA</i>	374	0,7%	227	31,3%	188,6
<i>PERCENTUALE</i>	356	0,6%	232	32,0%	176,2
<i>GLI</i>	346	0,6%	229	31,6%	173,2
<i>NEL</i>	345	0,6%	220	30,3%	178,7
<i>AL</i>	339	0,6%	220	30,3%	175,6
<i>NELLE</i>	336	0,6%	215	29,7%	177,4
<i>ANNO</i>	331	0,6%	228	31,5%	166,3
<i>DELLE</i>	330	0,6%	234	32,3%	162,1

Fonte: elaborazione INVALSI, dati RAV a.s. 2014/2015

Tabella 2.10 - Parole a frequenza media presenti nei punti di debolezza dell'area Successo scolastico

	FREQ.	% FREQ.	N. CASI	% CASI	TF • IDF
CLASSI	313	0,6%	195	26,9%	178,5
DEI	313	0,6%	219	30,2%	162,7
CLASSE	303	0,5%	204	28,1%	166,9
L	303	0,5%	209	28,8%	163,7
MEDIA	301	0,5%	190	26,2%	175,1
ALL	274	0,5%	218	30,1%	143
DELL	270	0,5%	195	26,9%	154
RISPETTO	263	0,5%	178	24,6%	160,4
SECONDARIA	226	0,4%	172	23,7%	141,2
NAZIONALE	225	0,4%	155	21,4%	150,8
FASCIA	219	0,4%	144	19,9%	153,7
AMMESSI	216	0,4%	163	22,5%	140
DA	216	0,4%	166	22,9%	138,3
VALUTAZIONE	213	0,4%	160	22,1%	139,8
NUMERO	203	0,4%	148	20,4%	140,1
ESAME	199	0,4%	173	23,9%	123,8
CORSO	193	0,3%	156	21,5%	128,8
STATO	187	0,3%	170	23,5%	117,8
VOTO	176	0,3%	134	18,5%	129
PRIMO	176	0,3%	148	20,4%	121,5
DATI	171	0,3%	132	18,2%	126,5
USCITA	165	0,3%	146	20,1%	114,8
ALLE	163	0,3%	130	17,9%	121,7
6	160	0,3%	124	17,1%	122,7
AI	154	0,3%	121	16,7%	119,7
SCOLASTICO	152	0,3%	123	17,0%	117,1
TRA	150	0,3%	119	16,4%	117,7
TRASFERIMENTI	149	0,3%	121	16,7%	115,9
ANCHE	146	0,3%	121	16,7%	113,5
ISTITUTO	144	0,3%	114	15,7%	115,7
1	143	0,3%	98	13,5%	124,3
PRIMARIA	142	0,3%	113	15,6%	114,6
GRADO	140	0,3%	118	16,3%	110,4
DATO	135	0,2%	93	12,8%	120,4
PARTE	135	0,2%	114	15,7%	108,5
SUCCESSIVA	134	0,2%	115	15,9%	107,2
7	133	0,2%	97	13,4%	116,2
NELL	132	0,2%	104	14,3%	111,3
	132	0,2%	108	14,9%	109,2
D	127	0,2%	113	15,6%	102,5
2	126	0,2%	78	10,8%	122
AD	126	0,2%	99	13,7%	109
SUPERIORE	126	0,2%	100	13,8%	108,4
BIENNIO	125	0,2%	83	11,5%	117,7
DISCIPLINE	122	0,2%	104	14,3%	102,9
0	121	0,2%	104	14,3%	102

FASCE	119	0,2%	93	12,8%	106,1
DIFFICOLTÀ	119	0,2%	99	13,7%	102,9
FAMIGLIE	119	0,2%	107	14,8%	98,9
GIUDIZIO	113	0,2%	96	13,2%	99,2
3	112	0,2%	74	10,2%	111
SOPRATTUTTO	111	0,2%	97	13,4%	97
RISULTA	110	0,2%	88	12,1%	100,7
ALTA	109	0,2%	93	12,8%	97,2
ED	107	0,2%	89	12,3%	97,5
10	105	0,2%	78	10,8%	101,7
DALLA	104	0,2%	89	12,3%	94,7
HANNO	103	0,2%	79	10,9%	99,2
SIA	101	0,2%	66	9,1%	105,1
AGLI	101	0,2%	88	12,1%	92,5
BASSA	99	0,2%	86	11,9%	91,7
RISULTATI	98	0,2%	85	11,7%	91,2
5	95	0,2%	60	8,3%	102,8
8	94	0,2%	77	10,6%	91,5
MEDIE	93	0,2%	72	9,9%	93,3
DEBITI	92	0,2%	79	10,9%	88,6
REGIONALE	91	0,2%	70	9,7%	92,4
DISTRIBUZIONE	91	0,2%	88	12,1%	83,3
INFERIORE	90	0,2%	78	10,8%	87,1
LICEO	89	0,2%	46	6,3%	106,6
4	89	0,2%	53	7,3%	101,1
RECUPERO	89	0,2%	74	10,2%	88,2
CRITERI	89	0,2%	82	11,3%	84,2
INDIRIZZO	88	0,2%	63	8,7%	93,4
DAL	87	0,2%	80	11,0%	83,3
NAZIONALI	85	0,2%	72	9,9%	85,3
RIFERIMENTO	84	0,2%	57	7,9%	92,8
MAGGIORE	84	0,2%	69	9,5%	85,8
ESITI	83	0,2%	68	9,4%	85,3
PERCENTUALI	82	0,1%	54	7,5%	92,5
CASI	81	0,1%	66	9,1%	84,3
FORMATIVI	81	0,1%	75	10,3%	79,8
COMPETENZE	80	0,1%	68	9,4%	82,2
ANNI	79	0,1%	73	10,1%	78,8
9	78	0,1%	62	8,6%	83,3
HA	78	0,1%	69	9,5%	79,7
MEDIO	77	0,1%	65	9,0%	80,7
PRIME	77	0,1%	66	9,1%	80,1
QUANTO	77	0,1%	67	9,2%	79,6
MA	77	0,1%	68	9,4%	79,1
TECNICO	76	0,1%	44	6,1%	92,5
VALUTAZIONI	75	0,1%	55	7,6%	84
PRIMA	75	0,1%	59	8,1%	81,7

<i>NEI</i>	75	0,1%	61	8,4%	80,6
<i>ESSERE</i>	74	0,1%	62	8,6%	79
<i>FORMATIVO</i>	74	0,1%	67	9,2%	76,5
<i>ALCUNI</i>	74	0,1%	67	9,2%	76,5
<i>ABBANDONI</i>	73	0,1%	68	9,4%	75
<i>CONCENTRANO</i>	73	0,1%	70	9,7%	74,1
<i>SEMPRE</i>	72	0,1%	62	8,6%	76,9
<i>LIVELLO</i>	72	0,1%	62	8,6%	76,9
<i>PARTICOLARE</i>	72	0,1%	63	8,7%	76,4
<i>ABBANDONO</i>	72	0,1%	68	9,4%	74
<i>SOCIO</i>	71	0,1%	64	8,8%	74,8
<i>VOTAZIONE</i>	70	0,1%	55	7,6%	78,4
<i>STRANIERI</i>	68	0,1%	52	7,2%	77,8
<i>PUNTI</i>	68	0,1%	53	7,3%	77,3
<i>CONSEGUITO</i>	68	0,1%	66	9,1%	70,8
<i>DOCENTI</i>	67	0,1%	60	8,3%	72,5
<i>VOTI</i>	65	0,1%	47	6,5%	77,2
<i>CONCENTRAZIONE</i>	65	0,1%	54	7,5%	73,3

Fonte: elaborazione INVALSI, dati RAV a.s. 2014/2015