

# LE PAROLE DEL RAV

## L'analisi testuale per la metavalutazione

Analisi delle motivazioni al giudizio:  
area dei Risultati nelle Prove Standardizzate  
Nazionali





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

Isabella Quadrelli 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 Prantero, Isabella Quadrelli (consulente), Maria Ranieri (consulente), Sara Romiti, Simone Russo, Stefania Sette, Consuelo Torelli (assegnista), Emanuela Vinci.

## 5.1 Introduzione

In questo capitolo sono illustrati i risultati dell'analisi testuale condotta con l'ausilio del software WordStat sul campo aperto Motivazioni del giudizio del Rapporto di Autovalutazione (RAV) nell'area Risultati nelle prove standardizzate nazionali del RAV.

In quest'area le scuole sono chiamate a riflettere sul livello di competenze raggiunto in Italiano e Matematica alle Prove INVALSI dagli studenti in relazione con i risultati raggiunti dalle altre scuole del territorio, nonché con quelli di scuole con background socio-economico simile. Inoltre, le scuole possono valutare la loro capacità di assicurare a tutti gli studenti il raggiungimento di livelli di apprendimento soddisfacenti, attraverso la riduzione dell'incidenza numerica e della dimensione del *gap* formativo degli studenti collocati nei livelli più bassi di apprendimento e limitando la variabilità di risultati interna alla scuola (tra le classi, tra le sedi, tra gli indirizzi). Nell'anno scolastico 2014-2015, nel quale sono stati redatti i RAV oggetto della presente analisi, non erano disponibili misure di valore aggiunto delle scuole.

Nel capitolo sono illustrati i risultati dell'analisi lessicale che ha preso in considerazione le parole a frequenza elevata e media, le parole specifiche dell'area e le parole comuni e originali (cfr. cap. 2). Inoltre saranno presentati i temi individuati con la procedura *Topic extraction* e il vocabolario tematico dell'area.

Gli obiettivi dell'analisi sono di valutare se gli strumenti di autovalutazione proposti alle scuole sono risultati efficaci nell'orientare le istituzioni scolastiche verso un percorso di autovalutazione *data driven*, e coerente, e se le scuole hanno adottato un orientamento riflessivo e articolato nell'analisi dei dati.

## 5.2 Analisi lessicale

Il corpus testuale della partizione Motivazioni del giudizio dell'area Risultati nelle prove standardizzate nazionali è composto da 67.020 *tokens* e da 3.278 *types*. In media ogni documento è composto da circa 100 parole (98). La partizione in esame risulta di dimensioni sufficienti per effettuare un'analisi di tipo statistico. Il quoziente *type/token* è pari a 0,052<sup>1</sup> e la quota di *hapax* risulta inferiore al 50%.

Tra le forme grafiche a elevata frequenza<sup>2</sup> troviamo parole congruenti con l'ambito semantico specifico dell'area in esame. Tra le parole piene<sup>3</sup>, troviamo il sostantivo "Matematica" (= 1480), seguito da "Italiano" (= 1450) (Fig. 5.1). È evidente il riferimento al contenuto delle prove standardizzate nazionali. Risulta rilevante anche la trattazione di alcuni aspetti dell'analisi dei dati. Parole come "media" (= 1301), "superiore" (= 604), "inferiore" (= 515), "livelli" (= 531), rientrano nella fascia alta di frequenza e risultano anche parole specifiche di quest'area. Esse infatti sono sovrautilizzate nei testi delle motivazioni dell'area in esame rispetto

<sup>1</sup> Un rapporto type/token superiore al 20% caratterizza un corpus testuale non sufficientemente ricco dal punto di vista lessicale per effettuare un'analisi di tipo quantitativo (Bolasco 1999).

<sup>2</sup> Le fasce di frequenza sono state individuate utilizzando il criterio suggerito da Bolasco (1999). Partendo dal fondo della distribuzione di frequenze, ovvero dagli *apax*, la fascia bassa di frequenza termina

quando compare il primo salto nella sequenza delle frequenze; la fascia media termina e inizia quella alta quando si incontra l'ultima parola con la stessa frequenza di quella precedente.

<sup>3</sup> L'analisi è stata condotta applicando la stop list. Pertanto, sono state escluse le parole che hanno una funzione grammaticale e quelle ritenute non significative ai fini dell'analisi.

all'intero corpus costituito dalle motivazioni di tutte le aree.

**Figura 5.1 - Parole piene a frequenza elevata relative alle Motivazioni del giudizio dell'area Prove standardizzate nazionali**



Tra le parole a frequenza media troviamo “varianza” (= 404), “discostano” (= 381), “background” (= 410), “socio” (= 395), “economico” (= 378), “culturale” (= 385), “punteggio” (= 365), “dati” (= 214), “dato” (= 101), “percentuale” (= 96), presumibilmente utilizzate dalle scuole per illustrare la propria situazione con riferimento ai risultati raggiunti. I termini “primaria” (= 263), “secondaria” (= 184),

“grado” (= 98), “seconde” (= 115), “quinte” (= 104), “liceo” (= 66) rimandano invece a un’analisi dettagliata della situazione interna della scuola Le parole “regionale” (= 97), “Italia” (= 66), “sud” (= 58) fanno riferimento a un’analisi dei risultati che tiene in considerazione il confronto con i valori di riferimento territoriali (media regionale, di macroarea geografica e nazionale) (Fig. 5.2).

**Figura 5.2 - Parole piene a frequenza media relative alle Motivazioni del giudizio dell'area Prove standardizzate nazionali**



Il vocabolario dell'area presenta molti termini comuni a quelli presenti nel vocabolario dei descrittori della rubrica e a quello delle domande guida. Solo il 27,4% delle occorrenze del corpus è costituito da parole originali. Il 63% delle occorrenze è rappresentato da parole presenti anche nei descrittori e nel 10% circa da parole contenute nelle domande guida.

Per l'area in esame, alcune delle parole originali maggiormente frequenti rimandano all'analisi e alla descrizione dei risultati realizzati dalla scuola. Si tratta infatti delle parole: "primaria", "secondaria", "grado", "inferiori", "superiori", "dato", "percentuale".

Questa prima analisi del profilo lessicale del corpus testuale evidenzia che i termini più frequenti risultano peculiari dell'area e, tra questi, in buona parte si trovano parole non presenti nei descrittori della rubrica e nelle domande guida. Ciò permette di ipotizzare che, nonostante il ricorso a un lessico specialistico, veicolato anche attraverso lo strumento di autovalutazione, le scuole abbiano effettuato un certo grado di elaborazione delle motivazioni. Esse, infatti, sembrano concentrarsi su un'analisi articolata basata sui dati, in alcuni casi focalizzata sulle differenze interne alla

scuola e/o sulle differenze territoriali. Si tratta evidentemente di un tentativo di integrare e personalizzare le motivazioni del giudizio a partire dagli stimoli forniti dagli strumenti. In particolare, la frequenza delle parole che rimandano alla descrizione della situazione interna alla scuola risulta maggiormente in linea con gli stimoli forniti dalle domande guida (che invitavano a riflettere sulle differenze interne alla scuola) piuttosto che con i descrittori della rubrica nei quali viene considerata la scuola nel suo complesso come riferimento per il confronto tra i punteggi delle classi e quello medio dell'istituto scolastico. Inoltre, in alcuni casi, le scuole sembrano concentrarsi anche sul confronto territoriale subnazionale, ovvero regionale e di macroarea geografica. Anche in questo caso si tratta di un'integrazione operata dalle scuole rispetto agli stimoli della rubrica di valutazione che fa riferimento esclusivamente al confronto con la media nazionale.

### 5.3 I temi emergenti

La procedura *Topic extraction* di Wordstat è stata utilizzata per individuare alcuni nuclei tematici all'interno del corpus testuale. L'analisi delle componenti

principali, su cui si basa la *Topic extraction*, ha permesso di individuare otto temi principali (tab. 5.1)<sup>4</sup>.

**Tabella 5.1 - Temi estratti con la procedura *Topic extraction* nell'area Risultati prove standardizzate nazionali (Motivazione del giudizio)**

TOPIC	KEYWORDS	EIGENVALUE	% VAR	FREQ.	N. CASI	% CASI
<b>BACKGROUND SOCIO-ECONOMICO</b>	ECONOMICO; SOCIO; SIMILE; BACKGROUND; CULTURALE; SCUOLE; PUNTEGGIO; INVALSI; PROVE; SCUOLA; SUPERIORE; ITALIANO; MATEMATICA	18,94	8,13	782	505	69,9%
<b>ANALISI DATI</b>	ISTITUTO; DATI; RISULTATI; ANNI; ANALISI; COMPETENZE; ESITI	6,08	3,45	1027	430	59,5%
<b>LIVELLI DI ABILITA' E VARIANZA</b>	QUOTA; COLLOCATA; 1; 2; STUDENTI; LIVELLI; NAZIONALE; INFERIORE; MEDIA; MATEMATICA; ITALIANO; DECISAMENTE; SUPERIORE; VARIANZA; TRA	3,57	7,65	4946	610	84,4%
<b>DESCRITTORE LIVELLO 3</b>	SINGOLE; NEGATIVO; CASI; MOLTO; ANCHE; DISCOSTANO; UGUALE; PUNTEGGI; POCO; NON	3,19	4,64	1961	522	72,2%
<b>CONFRONTO INTERNO SCUOLA</b>	GRADO; SECONDARIA; PRIMO; PRIMARIA SECONDE; QUINTE	2,52	2,27	316	198	27,4%
<b>CONFRONTO TERRITORIALE</b>	NAZIONALI; MEDIE; REGIONALI; SUPERIORI; RISULTATI; STANDARDIZZATE ITALIA; SUD	2,03	2,85	437	225	31,1%
<b>ELEVATI LIVELLI DI ABILITA'</b>	5; 4	1,69	1,85	102	87	12%
<b>VARIABILITA'</b>	ALTA; VARIABILITA'; PIÙ; ALTO	1,63	2,47	303	192	26,6%

FONTE: elaborazioni INVALSI dati RAV a.s. 2014/2015

La componente denominata **Background socio-economico** richiama il confronto tra i risultati ottenuti dalla scuola nelle prove standardizzate nazionali di Italiano e Matematica con i risultati di scuole con background socio-economico-culturale simile. Si tratta del primo aspetto, in ordine di sequenza, introdotto dai descrittori della rubrica. Le parole che co-occorrono nell'ambito della stessa frase sono quelle presenti nei descrittori. Questa componente spiega una quota elevata della varianza totale (8,13) e le parole chiave a

essa associate sono distribuite nel 70% circa dei documenti. L'analisi dei segmenti ripetuti (tab. 5.2) conferma che le parole chiave associate a questo tema co-occorrono molto frequentemente tra di loro. Inoltre, segmenti lunghi, con 5 o più parole tendono a essere meno diffusi rispetto a segmenti più corti. Fanno eccezione alcune espressioni inserite nei descrittori e ormai diventate di uso comune come "background socio economico e culturale" diffusa in oltre il 45% dei documenti.

<sup>4</sup> Peso fattoriale minimo pari a 0,40.

**Tabella 5.2 - Frequenza e diffusione dei segmenti ripetuti coerenti con il topic *Background socio economico***

	FREQ.	N. CASI	% CASI	LUNGHEZZA
ITALIANO E MATEMATICA	618	357	49,2%	3
PROVE INVALSI	423	388	53,5%	2
SOCIO ECONOMICO	375	360	49,7%	2
BACKGROUND SOCIO ECONOMICO	356	347	47,8%	3
SCUOLE CON BACKGROUND	355	344	47,5%	3
SOCIO ECONOMICO E CULTURALE	351	339	46,7%	4
CULTURALE SIMILE	350	342	47,2%	2
BACKGROUND SOCIO ECONOMICO E CULTURALE	339	330	45,5%	5
SCUOLE CON BACKGROUND SOCIO ECONOMICO	334	328	45,2%	5
BACKGROUND SOCIO ECONOMICO E CULTURALE SIMILE	329	323	44,6%	6
SCUOLE CON BACKGROUND SOCIO ECONOMICO E CULTURALE SIMILE	317	311	42,9%	8
*MATEMATICA DELLA SCUOLA	251	249	34,3%	3
PUNTEGGIO DI ITALIANO E MATEMATICA	245	243	33,5%	5
SCUOLA ALLE PROVE INVALSI	245	242	33,4%	4
MATEMATICA DELLA SCUOLA ALLE PROVE INVALSI	226	225	31%	6
ITALIANO E MATEMATICA DELLA SCUOLA	215	213	29,4%	5
PUNTEGGIO DI ITALIANO E MATEMATICA DELLA SCUOLA	203	202	27,9%	7
ITALIANO E MATEMATICA DELLA SCUOLA ALLE PROVE INVALSI	195	194	26,8%	8
PUNTEGGIO DI ITALIANO E MATEMATICA DELLA SCUOLA ALLE PROVE	189	188	25,9%	9
SUPERIORE A QUELLO DI SCUOLE CON BACKGROUND SOCIO ECONOMICO	135	135	18,6%	9

FONTE: elaborazioni INVALSI dati RAV a.s. 2014/2015

Allo stesso modo, il tema denominato **Livelli di abilità e varianza** spiega una buona quota di varianza totale (7,65). A esso sono associate le parole chiave che rimandano all'analisi dei livelli di abilità degli studenti (soprattutto dei livelli 1 e 2) e della varianza tra e dentro le classi. Anche in questo caso si tratta di temi proposti dai descrittori e ripresi dalle scuole che vi fanno riferimento all'interno della stessa proposizione. L'analisi dei segmenti ripetuti (tab. 5.3) mette in evidenza come le scuole utilizzino con una certa frequenza espressioni di uso comune come 'media nazionale', 'varianza tra' presenti anche nei descrittori, oppure riprendano alcune espressioni dei descrittori come 'livelli 1 e 2' per illustrare la propria situazione. La diffusione nei documenti di segmenti più lunghi, che riproducono fedelmente il testo dei descrittori, è minore e in genere riguarda poco più o poco meno del 30% dei documenti (es. "quota di studenti collocata nei livelli 1 e 2", "classi in Italiano e Matematica").

Un'altra componente individua la cooccorrenza nella stessa frase di parole che

possono essere ricondotte al descrittore del livello 3 della rubrica nel quale si fa riferimento al confronto tra i risultati delle singole classi e quello della scuola nel suo complesso. In un buon numero di casi, le scuole che si sono collocate tra i livelli 2 e 4 nella scala di valutazione hanno riproposto o adattato il testo di una parte del descrittore del livello 3, ovvero:

*I punteggi delle classi non si discostano molto dalla media della scuola, anche se ci sono casi di singole classi in Italiano e Matematica che si discostano in negativo.*

Tra le parole che compongono questo tema, la presenza di aggettivi come "uguale" e "poco", non presenti nel testo originale del descrittore, fanno pensare a una rielaborazione da parte delle scuole che effettuano degli adattamenti degli input proposti dagli strumenti operando delle variazioni su tema. Le scuole, ad esempio, inseriscono degli aggettivi nella struttura originaria dei segmenti dei descrittori che hanno la funzione di caratterizzare in modo più adeguato la situazione che vogliono descrivere e rappresentare.

**Tabella 5.3 - Frequenza e diffusione dei segmenti ripetuti coerenti con il tema “Livelli di qualità e varianza”**

	FREQ.	N. CASI	% CASI	LUNGHEZZA
MEDIA NAZIONALE	577	388	53,5%	2
CLASSI IN ITALIANO	378	272	37,2%	3
VARIANZA TRA	336	320	44,1%	2
1 E 2	334	312	43%	3
ITALIANO E IN MATEMATICA	328	244	33,7%	4
QUOTA DI STUDENTI	322	296	40,8%	3
STUDENTI COLLOCATA	320	293	40,4%	2
LIVELLI 1	317	302	41,7%	2
QUOTA DI STUDENTI COLLOCATA	300	277	38,2%	4
LIVELLI 1 E 2	299	284	39,2%	4
TRA CLASSI	281	275	37,9%	2
STUDENTI COLLOCATA NEI LIVELLI	280	267	36,8%	4
QUOTA DI STUDENTI COLLOCATA NEI LIVELLI	266	255	35,2%	6
STUDENTI COLLOCATA NEI LIVELLI 1	266	259	35,7%	5
2 IN ITALIANO	263	257	35,5%	3
VARIANZA TRA CLASSI	261	257	35,5%	3
STUDENTI COLLOCATA NEI LIVELLI 1 E 2	254	247	34,1%	7
QUOTA DI STUDENTI COLLOCATA NEI LIVELLI 1	252	247	34,1%	7
LIVELLI 1 E 2 IN ITALIANO	244	238	32,8%	6
QUOTA DI STUDENTI COLLOCATA NEI LIVELLI 1 E 2	241	236	32,6%	9
CLASSI IN ITALIANO E MATEMATICA	239	197	27,2%	5
STUDENTI COLLOCATA NEI LIVELLI 1 E 2 IN ITALIANO	232	227	31,3%	9
VARIANZA TRA CLASSI IN ITALIANO	228	226	31,2%	5
INFERIORE ALLA MEDIA	219	203	28%	3
INFERIORE ALLA MEDIA NAZIONALE	200	189	26,1%	4
LIVELLI 1 E 2 IN ITALIANO E IN MATEMATICA	196	192	26,5%	9

FONTE: elaborazioni INVALSI dati RAV a.s. 2014/2015

Il tema denominato **Analisi dati** è stato individuato a partire dall'associazione con alcune parole chiave che rimandano principalmente alla presentazione e discussione dei risultati o dei dati della scuola (“dati”, “risultati”, “anni”, “analisi”, “esiti”). Si tratta di parole che non sono presenti nei descrittori della rubrica. Il sostantivo “risultati” compare nelle domande guida oltre che nel titolo dell'area. Nei testi nei quali ricorrono queste parole l'analisi dei dati viene utilizzata per esprimere un giudizio sulla capacità della scuola di assicurare l'acquisizione delle competenze. Proponiamo alcuni passaggi a titolo di esempio.

*Il livello di istituto è globalmente superiore alla media nazionale e regionale in tutte le prove, particolare eccellenza per i punteggi*

*relativi a Italiano. I dati rilevati indicano quindi l'efficacia dell'azione messa in atto dalla nostra scuola per lo sviluppo delle competenze specifiche e trasversali testate dall'INVALSI, frutto di un lavoro di più anni di lettura, interpretazione, confronto dei dati. [caso 56]*

*Il livello raggiunto dalla maggior parte delle classi in Italiano e Matematica è positivo, la scuola assicura l'acquisizione dei livelli essenziali di competenze. [caso 63]*

*I risultati delle prove standardizzate nazionali evidenziano alcune criticità, riconducibili almeno in parte alla scarsa partecipazione degli studenti alla rilevazione. I dati a disposizione sono parziali, non consentono un'analisi esaustiva sul livello di competenze raggiunto in entrambe le discipline nei diversi gradi di istruzione né di rilevare la*

	<p><i>variabilità dei risultati tra le classi e al loro interno. Gli esiti di apprendimento di una sola classe, se non addirittura fuorvianti, non sono sufficienti per misurare il livello di acquisizione delle competenze di Italiano nella scuola primaria e per rilevare i risultati conseguiti a livello d'istituto. [caso 101]</i></p> <p><i>I risultati positivi sono la conseguenza di una politica di insegnamento che tende all'acquisizione di competenze e non solo di conoscenze. La varianza tra le classi è in alcuni casi ancora sensibile, per cui non si è riusciti a contrastare del tutto gli effetti di discrepanza. Si ravvisa l'esigenza di incrementare il lavoro sulle competenze logico-matematiche e di consolidare quello svolto nell'area delle competenze linguistiche in L1. Le nuove modalità di formazione delle classi hanno permesso di correggere alcuni effetti di disparità, per cui i risultati nelle seconde sono più omogenei rispetto ai risultati nelle quinte. [caso 118]</i></p> <p><i>Sulla base della lettura e dell'analisi dei dati riportati si evince che l'istituto assicura nel complesso una sufficiente acquisizione dei livelli essenziali di competenze in Italiano, ma in misura non accettabile per quanto concerne la Matematica. [caso 139]</i></p>
--	---

Si tratta di testi nei quali le scuole presentano o discutono i risultati del proprio lavoro e nei quali è presente un certo grado di rielaborazione del discorso rispetto al testo dei descrittori.

La diffusione delle parole chiave associate a questo tema in una buona percentuale di casi (circa il 60%) fa pensare che sia piuttosto frequente la pratica di inserire nel testo della motivazione alcuni passaggi nei quali le scuole operano degli approfondimenti, descrivendo la propria situazione con riferimento ai dati, oppure esprimono un giudizio contestualizzato.

Altre co-occorrenze tra le parole chiave mettono in evidenza l'emergere di temi che rimandano all'articolazione dell'analisi dei dati. Si tratta del tema **Confronto interno scuola** definito a partire dalla co-occorrenza di parole che

evidenziano una descrizione dei risultati disaggregata per ordine di scuola o a un confronto tra classi della stessa scuola. Le parole chiave associate sono infatti: "grado", "secondaria", "primo", "primaria", "seconde", "quinte". Il tema **Confronto territoriale** invece è associato a parole chiave che richiamano un confronto con i valori di riferimento territoriali ("nazionali", "regionali", "Italia", "sud"). Questi temi sono diffusi rispettivamente nel 27,4% e nel 31,1% dei casi.

Infine, gli ultimi due temi riguardano rispettivamente gli studenti con livelli elevati di abilità (**Elevati livelli di abilità**) - presente soprattutto nelle motivazioni delle scuole che si sono assegnate punteggi alti (6 e 7) - e la **Varianza tra le classi**. La componente riferita a quest'ultimo tema, tuttavia, intercetta le proposizioni nelle quali le scuole utilizzano il sostantivo 'variabilità' e non 'varianza', come indicato nei descrittori della rubrica.

L'analisi delle componenti principali ha evidenziato la presenza nel corpus testuale di nuclei tematici riconducibili sia agli specifici aspetti proposti dai descrittori della rubrica – il confronto con scuole con background socio-economico e culturale simile, i livelli di abilità degli studenti, la varianza tra le classi e il confronto tra i risultati delle singole classi e quello della scuola – sia alle riflessioni proposte con le domande guida – l'analisi disaggregata a livello di scuola. Si possono individuare inoltre nuclei tematici riconducibili a contesti discorsivi nei quali vengono illustrati i dati o discussi i risultati della scuola e, in una minoranza di casi, proposizioni nelle quali le scuole dedicano spazio al confronto territoriale nelle sue articolazioni subnazionali. Quest'ultimo aspetto rappresenta una prospettiva introdotta autonomamente dalle scuole poiché non sollecitata né nelle domande guida, né nei descrittori della rubrica.

### 5.3 Il vocabolario tematico

A partire dalla struttura tematica evidenziata con l'analisi delle componenti principali è stato costruito un vocabolario tematico consolidando alcuni nuclei semantici mediante la selezione delle parole chiave maggiormente significative, la rimozione delle parole meno caratterizzanti perché utilizzate in una molteplicità di contesti e l'aggiunta di nuove parole chiave. In particolare, i temi evidenziati con la procedura *Topic extraction* sono stati trasformati in categorie tematiche. Ogni categoria rappresenta un nucleo di significato e i termini che vi sono inclusi rappresentano degli indicatori parziali del tema individuato. Ciascuna categoria è stata integrata con ulteriori termini ritenuti congruenti dal punto di vista semantico. Sono state aggiunte anche due categorie, una per i termini a valenza positiva e una per quelli a valenza negativa. L'inserimento di queste categorie ha lo scopo di individuare le parole utilizzate per qualificare le situazioni positive e negative e di valutare la congruenza tra la frequenza di tali forme grafiche e il punteggio autoattribuito nella scala di valutazione.

Il vocabolario tematico presenta un buon livello di copertura. Dopo aver applicato la *stop list*, esso copre il 76,1% delle frasi e il 99% dei documenti.

La tabella 5.4 presenta le categorie del vocabolario tematico, la loro frequenza in valori assoluti e percentuali, la diffusione in termini assoluti e percentuali delle categorie nei documenti e il valore dell'indice tf-idf. La tabella 5.5 presenta la frequenza e diffusione nei casi delle keyword associate a ciascuna categoria del vocabolario tematico.

**Tabella 5.4 - Frequenza e diffusione delle categorie del vocabolario tematico delle Motivazioni del giudizio (Risultati nelle prove standardizzate nazionali)**

	FREQ.	% FREQ.	N. CASI	% CASI	TF • IDF
BACKGROUND_SOCIO-ECONOMICO	5397	33,4%	627	86,7%	333,9
LIVELLI_ABILITA	3301	20,4%	563	77,9%	358,6
CONFRONTO_CLASSI-SCUOLA	2278	14,11	539	74,6%	290,6
RISULTATI_PROVE_POS	1276	7,9%	540	74,7%	161,7
ANALISI_DATI	1098	6,8%	410	56,7%	270,5
CONFRONTO_INTERNO_SCUOLA	881	5,5%	330	45,6%	300,1
RISULTATI_PROVE_NEG	880	5,5%	481	66,5%	155,8
CONFRONTO_TERRITORIALE	561	3,5%	226	31,3%	283,3
VARIANZA	477	3%	396	54,8%	124,7

FONTE: elaborazioni INVALSI dati RAV a.s. 2014/2015

La categoria **Background socio-economico** contiene le parole utilizzate per analizzare i risultati della scuola nelle prove standardizzate di Matematica e di Italiano in confronto con quelli di scuole con un simile background socio-economico-culturale. Le parole associate a questa categoria ricorrono con molta frequenza (33,2%) e sono diffuse nell'86,7% dei documenti analizzati.

La categoria **Livelli di abilità** comprende alcune parole chiave utilizzate per descrivere la situazione della scuola con riferimento alla distribuzione dei risultati degli studenti nelle prove standardizzate per livelli di abilità. Questa categoria comprende le parole associate all'omonimo tema individuato con la procedura *Topic extraction* e al tema "Elevati livelli di abilità". L'unione delle due categorie è avvenuta in seguito alla valutazione dell'omogeneità semantica dei due raggruppamenti. Le parole associate alla categoria evidenziano che le scuole prevalentemente descrivono la situazione dei propri studenti effettuando confronti tra la percentuale di studenti collocati nei livelli 1 e 2 e la media nazionale. In una minoranza di casi, le scuole si soffermano sull'analisi dei dati relativi agli studenti collocati nei livelli più alti (4 e 5). Questa categoria è diffusa nel 78% circa dei casi.

La categoria **Confronto classi-scuola** è costituita dalle parole chiave utilizzate per descrivere la situazione delle

singole classi con riferimento al punteggio medio della scuola. Questa categoria contiene la maggior parte delle parole associate al tema "Descrittore livello 3". Infatti, la gran parte delle parole chiave è comune al descrittore della rubrica riferito al livello 3. I termini associati a questa categoria sono diffusi in un'ampia parte del corpus testuale, ovvero nel 76,5% dei documenti.

La categoria **Varianza** si riferisce alle parole utilizzate per trattare il tema della varianza dei risultati degli studenti tra le classi. In virtù della loro omogeneità semantica, sono stati inclusi in questa categoria entrambi i sostantivi 'varianza' e 'variabilità' anche se collocati in contesti discorsivi diversi. L'analisi delle componenti principali aveva infatti individuato un nucleo semantico associato al termine 'varianza', utilizzato all'interno di una struttura discorsiva più vicina a quella dei descrittori della rubrica, e un secondo nucleo, associato al termine 'variabilità'. L'utilizzo di questo sinonimo, non presente nei descrittori, fa pensare a un certo grado di rielaborazione del testo delle motivazioni. I termini associati a questa categoria sono presenti nel 55% circa dei documenti.

Le parole chiave incluse nella categoria **Analisi dei dati** sono meno frequenti ma coprono complessivamente il 56,7% dei casi. Poiché le parole chiave associate a questa categoria si riferiscono

alla presentazione e discussione dei dati, si evince che in oltre una metà dei casi le scuole hanno articolato le motivazioni, facendo riferimento ai risultati raggiunti per giustificare il livello di qualità attribuito nella rubrica. L'ampia diffusione di termini che si riferiscono all'analisi dei dati indica che le scuole hanno effettuato una rielaborazione del testo dei descrittori e/o hanno integrato parti aggiuntive al fine di rendere conto della specificità della propria situazione.

Le categorie **Confronto interno scuola** e **Confronto territoriale** riprendono i temi individuati precedentemente con l'analisi delle componenti principali; le due categorie sono state consolidate con l'aggiunta di ulteriori parole chiave semanticamente congruenti.

La categoria "Confronto interno scuola" si caratterizza per la presenza di parole e segmenti come "secondaria", "secondaria di primo grado" che rimandano a un'articolazione dell'analisi dei dati che considera le differenze tra gli ordini di scuola. Parole quali "seconde", "quinte", ecc. fanno pensare a un'analisi disaggregata per classi all'interno di un medesimo ordine di scuola. Infine, altri termini rimandano a un'articolazione dell'analisi per sezioni e plessi e, nella secondaria di secondo grado, per indirizzo di scuola ("licei", "tecnico"). Le parole associate a questa categoria sono relativamente poche, tuttavia sono presenti nel 45,6% dei documenti analizzati.

La categoria che si riferisce all'analisi contestualizzata a livello territoriale è invece presente in poco più del 30% dei documenti.

Infine, le due categorie che raggruppano i termini a **valenza positiva** e

**negativa** sono presenti rispettivamente nel 74,7% e nel 66,5% dei casi.

Tra i termini a valenza positiva, i più frequenti sono gli aggettivi che evidenziano l'idea di superiorità ("superior\*"<sup>5</sup>), rispetto a valori medi e altri criteri di riferimento, presenti nel 55,7% dei documenti e di positività ("positiv\*"), diffusi nel 26,5% dei documenti. Il segmento 'non si discostano molto' considerato a valenza positiva, ripreso dal descrittore del livello 3, è diffuso in quasi il 14% delle motivazioni. I termini a valenza negativa più diffusi sono quelli che esprimono l'idea di inferiorità (diffusi in quasi il 43% dei documenti) e di negatività (nel 18,6% dei casi).

---

<sup>5</sup> L'utilizzo del metacarattere "\*" permette di cercare parole che hanno una parte comune, come "positiv", ma declinazioni diverse come negli aggettivi "positivo",

"positiva", "positivi", "positive" e suffissi come nell'avverbio "positivamente".

**Tabella 5.5 - Frequenza e diffusione delle keyword nelle categorie del vocabolario tematico delle Motivazioni del giudizio (Risultati nelle prove standardizzate nazionali)**

BACKGROUND SOCIO-ECONOMICO			CONFRONTO INTERNO SCUOLA		
	FREQ.	% CASI		FREQ.	% CASI
MATEMATICA	1480	69,9%	SECONDE	115	13,4%
ITALIANO	1450	69,7%	SECONDARIA	113	11,8%
PUNTEGGIO	425	52,3%	QUINTE	104	11,3%
PROVE_INVALSI	423	53,7%	LICEO	53	4,3%
SIMILE	393	51,5%	TERZE	47	4,8%
CULTURALE	385	50,8%	SECONDARIA_DI_PRIMO_GRADO	44	5,7%
BACKGROUND_SOCIO_ECONOMICO	356	48%	SECONDA	36	4,3%
PROVE	242	25,9%	PLESSI	35	4,2%
INVALSI	81	9,7%	GRADO	27	3%
BACKGROUND	54	5,4%	SECONDARIA_DI_I_GRADO	27	3,2%
ESCS	47	5,4%	SEZIONI	25	2,9%
SOCIO	39	5,1%	TECNICO	25	2,4%
ECONOMICO	22	2,9%	II	24	2,4%
LIVELLI DI ABILITÀ			V	23	2,8%
	FREQ.	% CASI	PRIMO	21	2,9%
MEDIA_NAZIONALE	577	53,7%	INDIRIZZI	20	2,5%
LIVELLI	531	55,5%	PLESSO	20	2,4%
MEDIA	493	50,5%	QUINTA	20	2,4%
QUOTA	369	44,8%	III	15	1,8%
1_E_2	336	43,4%	LICEI	15	1%
COLLOCATA	334	41,6%	BIENNIO	13	1,2%
5	149	14,4%	LICEO_SCIENTIFICO	13	1,7%
NAZIONALE	141	14,6%	TERZA	13	1,7%
2	139	12%	ORDINI_DI_SCUOLA	12	1,7%
1	119	12%	ISTITUTO_TECHNICO	11	1%
4	113	12%	CLASSICO	10	1,1%
CONFRONTO CLASSI SCUOLA			CONFRONTO TERRITORIALE		
	FREQ.	% CASI		FREQ.	% CASI
CLASSI	1427	73,6%	NAZIONALI	140	15,2%
VARIANZA	697	63%	REGIONALE	97	10,2%
PUNTEGGI	365	43,4%	MEDIE	88	8,6%
DISCOSTANO	282	31,3%	ITALIA	66	5,4%
MEDIA_DELLA_SCUOLA	231	30,6%	SUD	58	4,4%
CASI	134	16,9%	REGIONALI	41	4,6%
SINGOLE	100	13,6%	OVEST	18	1,7%
VARIANZA			EST	17	1,4%
TRA	544	59,2%	PUGLIA	14	0,7%
VARIANZA	404	49,7%	LOMBARDIA	12	1,2%
VARIABILITÀ	73	7,2%	REGIONE	10	1,2%
ANALISI DEI DATI					
	FREQ.	% CASI			
RISULTATI	463	39,7%			
DATI	214	20,6%			
ESITI	150	16%			
COMPETENZE	84	9,1%			
ANALISI	74	9%			
CHEATING	59	6,2%			
ANNI	54	6,9%			

RISULTATI PROVE POS		
	FREQ.	% CASI
SUPERIOR*	715	55,7%
POSITIV*	231	26,6%
NON_SI_DISCOSTANO_MOLTO	99	13,7%
MIGLIOR*	73	9,4%
OMOGEN*	53	6,8%
BUON*	44	5,4%
SOPRA	26	3%
ADEGUAT*	21	2,5%
SODDISFACENT*	14	1,8%

RISULTATI PROVE NEG		
	FREQ.	% CASI
INFERIORE	446	42,9%
NEGATIV*	145	18,6%
CRITICITÀ	93	10,6%
POCO_INFERIORE	69	9,2%
SOTTO	46	5,2%
DISPARITÀ	34	4%
DIFFICOLTÀ	30	3,7%
DISOMOGEN*	17	2,2%

FONTE: elaborazioni INVALSI dati RAV a.s. 2014/2015

## 5.4 La struttura delle motivazioni

L'analisi delle co-occorrenze tra le categorie del vocabolario tematico (indice di similarità costruito a partire dal coefficiente di Jaccard) evidenzia che, nell'ambito dello stesso documento, le scuole tendono a trattare tutte le tematiche introdotte dai descrittori della rubrica. Infatti, le keyword associate alla categoria maggiormente diffusa, ovvero "Background socio-economico" occorrono insieme a quelle della categoria "Livelli di abilità" nell'82,5% dei casi e con i termini a valenza positiva in oltre l'80% delle motivazioni. Anche le parole chiave della categoria "Confronto classi scuola" ricorrono insieme a quelle delle categorie "Background socio-economico" (nel 79,1% delle motivazioni) e "Livelli abilità" (77,2%). Il tema della varianza ricorre invece insieme alle altre categorie nel 60% circa dei casi.

I termini a valenza negativa occorrono insieme a quelli della categoria "Livelli abilità" nel 73,7% dei casi e con quelli delle categorie "Confronto classi scuola" e "Background socio economico" in percentuali di poco inferiori (rispettivamente 72,3% e 71,3%).

Per quanto riguarda le categorie relative all'analisi dei risultati, le keyword della categoria "Confronto interno scuola" occorrono frequentemente insieme a quelle delle categorie "Confronto classi scuola" (nel 48,8% dei casi) e "Analisi dati" (nel 47,4% dei casi). Le parole chiave incluse nella categoria "Confronto territoriale" vengono invece utilizzate con maggiore frequenza insieme a quelle delle categorie "Analisi dati" (38,6%) e "Confronto interno scuola" (35,9%) (tab. 5.6).

L'analisi delle co-occorrenze ha permesso di far emergere la struttura tematica delle motivazioni. Come si poteva intuire anche dall'ampia diffusione nei casi di alcune categorie del vocabolario tematiche, quasi tutte le scuole hanno fatto

riferimento, per giustificare il punteggio autoattribuito, ai principali temi introdotti dalla rubrica. L'aspetto trattato con minore frequenza è la varianza dei risultati tra le classi. In circa la metà dei casi le scuole hanno utilizzato parole che rimandano alla discussione di dati e/o risultati, illustrando in maniera dettagliata la situazione interna delle scuole e/o facendo confronti con i valori di riferimento territoriali. Negli altri casi, le scuole sembrano limitarsi a esprimere giudizi positivi e negativi seguendo l'impostazione proposta dai descrittori della rubrica.

**Tabella 5.6 - Indici di similarità delle categorie del vocabolario tematico**

	BACKGROUND_SOCIO-ECONOMICO	ANALISI_DAT	CONFRONTO_CLASSI-SCUOLA	CONFRONTO_INTERNO_SCUOLA	CONFRONTO_TERRITORIALE	LIVELLI_ABILITA	RISULTATI_PROVE_NEG	RISULTATI_PROVE_POS	VARIANZA
BACKGROUND_SOCIO-ECONOMICO	1.000	0.564	0.791	0.491	0.339	0.825	0.713	0.809	0.606
ANALISI_DATI	0.564	1.000	0.480	0.474	0.386	0.490	0.421	0.489	0.306
CONFRONTO_CLASSI-SCUOLA	0.791	0.480	1.000	0.488	0.317	0.772	0.723	0.760	0.670
CONFRONTO_INTERNAL	0.491	0.474	0.488	1.000	0.359	0.459	0.451	0.462	0.380
CONFRONTO_TERRITORIALE	0.339	0.386	0.317	0.359	1.000	0.331	0.304	0.325	0.242
VELLI_ABILITA	0.825	0.490	0.772	0.459	0.331	1.000	0.737	0.791	0.645
RISULTATI_PROVE_NEG	0.713	0.421	0.723	0.451	0.304	0.737	1.000	0.699	0.652
RISULTATI_PROVE_POS	0.809	0.489	0.760	0.462	0.325	0.791	0.699	1.000	0.645
VARIANZA	0.606	0.306	0.670	0.380	0.242	0.645	0.652	0.645	1.000

FONTE: elaborazioni INVALSI dati RAV a.s. 2014/2015

## 5.5 Coerenza tra motivazioni e punteggi autoattribuiti

La compilazione del campo aperto “Motivazioni” nell’ambito del RAV è finalizzata a rendere conto delle ragioni che hanno portato le scuole a collocarsi in uno specifico livello nella scala di valutazione in relazione ad un criterio di qualità.<sup>6</sup>

Le scuole del campione si sono collocate su tutti i livelli della scala di autovalutazione; la distribuzione delle frequenze evidenzia una curva simmetrica. La maggior parte delle scuole, il 30% circa, si è collocata nel livello 4, ovvero il livello di qualità intermedio. Le altre scuole si sono distribuite simmetricamente nei livelli medio bassi e medio alti (tab. 5.6).

<sup>6</sup> Il criterio di qualità dell’area Risultati nelle prove standardizzate nazionali è “La scuola assicura l’acquisizione dei livelli essenziali di competenze (misurate con le prove standardizzate nazionali) per tutti gli studenti”. Il livello 1 della rubrica che identifica una situazione “molto critica” è così descritto: “Il punteggio di italiano e/o matematica della scuola alle prove INVALSI è inferiore rispetto a quello di scuole con background socio-economico e culturale simile. I punteggi delle diverse classi in italiano e/o matematica sono molto distanti e la varianza tra classi in italiano e/o matematica è decisamente superiore a quella media. La quota di studenti collocata nei livelli 1 e 2 in italiano

e in matematica è notevolmente superiore alla media nazionale”. Il livello 7, eccellente, è descritto nei seguenti termini: “Il punteggio di italiano e matematica della scuola alle prove INVALSI è superiore a quello di scuole con background socioeconomico e culturale simile ed è superiore alla media nazionale. La varianza tra classi in italiano e matematica è inferiore a quella media. I punteggi delle diverse classi in italiano e matematica non si discostano dalla media della scuola. La quota di studenti collocata nei livelli 1 e 2 in italiano e in matematica è decisamente inferiore alla media nazionale.

**Tabella 5.6 - Scuole collocate nei livelli della scala di Valutazione (Risultati nelle prove standardizzate nazionali)**

LIVELLI DELLA SCALA	FREQ.	% FREQ.
1	13	1.8%
2	85	11.8%
3	146	20.2%
4	216	29.9%
5	148	20.5%
6	90	12.4%
7	25	3.5%
<b>TOTALE</b>	<b>723</b>	<b>100%</b>

FONTE: elaborazioni INVALSI dati RAV a.s. 2014/2015

L'analisi della distribuzione delle categorie del vocabolario tematico per il punteggio autoattribuito (tab. 5.7) evidenzia, in primo luogo, che la distribuzione delle categorie che includono parole a valenza positiva e negativa è coerente con il posizionamento delle scuole nella scala di autovalutazione.

La frequenza delle keyword della categoria "Background socio-economico" è leggermente superiore nelle motivazioni delle scuole collocate nei livelli medio-bassi della scala di valutazione; d'altro canto, quelle associate alla categoria "Livelli di abilità" sono più frequenti nei documenti prodotti dalle scuole posizionate nei livelli più elevati. Le keyword della categoria

"Confronto territoriale" sono significativamente più frequenti nelle motivazioni delle scuole collocate nel livello più basso. Le keyword che fanno riferimento al tema della varianza sono invece più frequenti nelle motivazioni delle scuole collocate nei livelli 2, 4 e 6, che rappresentano i livelli privi di descrittore. Possiamo ipotizzare che le scuole autocollocate in questi livelli abbiano utilizzato più frequentemente le keyword associate alla categoria "Varianza" per illustrare con maggiore dettaglio la propria situazione rispetto a questo specifico indicatore. Non emergono differenze significative nell'utilizzo di parole che fanno riferimento all'analisi dei dati e al confronto interno alla scuola.

**Tabella 5.7 - Distribuzione delle categorie del vocabolario tematico per punteggio autoattribuito**

	1	2	3	4	5	6	7	Chi2	P (2-tails)
<b>BACKGROUND_SOCIO-ECONOMICO</b>	16,22%	15,73%	15,51%	15,05%	15,37%	14,68%	15,15%	23,61	0,001
<b>ANALISI_DATI</b>	9,46%	8,27%	11,56%	9,40%	11,43%	8,29%	12,12%	4,509	0,608
<b>CONFRONTO_CLASSI-SCUOLA</b>	12,16%	13,51%	13,01%	12,99%	13,67%	12,61%	12,88%	10,51	0,105
<b>CONFRONTO_INTERNO_SCUOLA</b>	6,76%	7,06%	8,94%	8,63%	7,62%	7,25%	6,82%	8,34	0,214
<b>CONFRONTO_TERRITORIALE</b>	10,81%	3,43%	6,83%	5,42%	5,26%	5,35%	5,30%	13,895	0,031
<b>LIVELLI_ABILITA</b>	12,16%	14,31%	13,14%	13,60%	13,80%	14,34%	12,88%	28,461	0
<b>RISULTATI_PROVE_NEG</b>	14,86%	12,90%	11,17%	11,99%	10,12%	12,44%	11,36%	34,836	0
<b>RISULTATI_PROVE_POS</b>	9,46%	13,31%	11,56%	12,91%	14,32%	14,34%	13,64%	35,685	0
<b>VARIANZA</b>	8,11%	11,49%	8,28%	10,01%	8,41%	10,71%	9,85%	31,799	0

FONTE: elaborazioni INVALSI dati RAV a.s. 2014/2015

L'analisi delle corrispondenze lessicali fornisce ulteriori elementi per qualificare il comportamento delle scuole in relazione alla stesura delle motivazioni (Fig. 5.3). L'ACL conferma in primo luogo come le motivazioni delle scuole posizionate sui livelli diversi della scala di autovalutazione si caratterizzino per l'utilizzo di alcune parole specifiche. Sul primo asse, nel versante positivo, si collocano le scuole che si sono attribuite il punteggio 3. Le motivazioni di queste scuole si caratterizzano per l'utilizzo di parole che esprimono una valenza negativa (disparità, difficoltà, criticità, negatività) ma anche per la presenza di termini che rimandano all'idea di adeguatezza. Ciò risulta coerente con il posizionamento nel livello medio-basso della scala. La diffusione di alcune parole come 'seconde', 'terze' e 'plesso' fanno pensare a un'analisi articolata a livello di scuola. Nel semiasse negativo, troviamo le motivazioni delle scuole che si sono collocate nei livelli 6 e 7. Le motivazioni di queste scuole si caratterizzano per l'utilizzo di termini presenti nei descrittori dei livelli 5 e 7. Si tratta della voce verbale 'collocata' utilizzata per illustrare la situazione degli studenti per livelli di abilità, e dell'aggettivo 'inferiore', impiegato in questo contesto per esprimere una condizione positiva, ovvero la presenza di una percentuale inferiore alla media nazionale di studenti collocati nei livelli 1 e 2. La vicinanza con i termini numerici '4' e '5' fa pensare al fatto che

queste scuole abbiano fatto riferimento anche alla percentuale di studenti collocati nei livelli più elevati per evidenziare, presumibilmente, la loro presenza in percentuale superiore rispetto al riferimento nazionale. La presenza del sostantivo 'liceo' rimanda a un'articolazione dell'analisi per indirizzo di scuola effettuata dagli istituti secondari.

Nel versante positivo del secondo asse si trovano le scuole posizionate nel livello 5 della scala di autovalutazione. Le loro motivazioni si caratterizzano per l'utilizzo di termini che esprimono positività ("sopra", "buon", "positivo") e per la presenza di parole che rimandano all'analisi dei dati ("dati") e al confronto all'interno della scuola ("sezioni"). Nel semiasse negativo sono posizionate le scuole che si sono autoattribuite i punteggi 2 e 4. I contributi più rilevanti in questo semiasse sono dati dai termini "casi" e "singoli" presenti nel descrittore del livello 3 e utilizzati per illustrare il confronto tra i risultati delle singole classi e il punteggio complessivo della scuola.

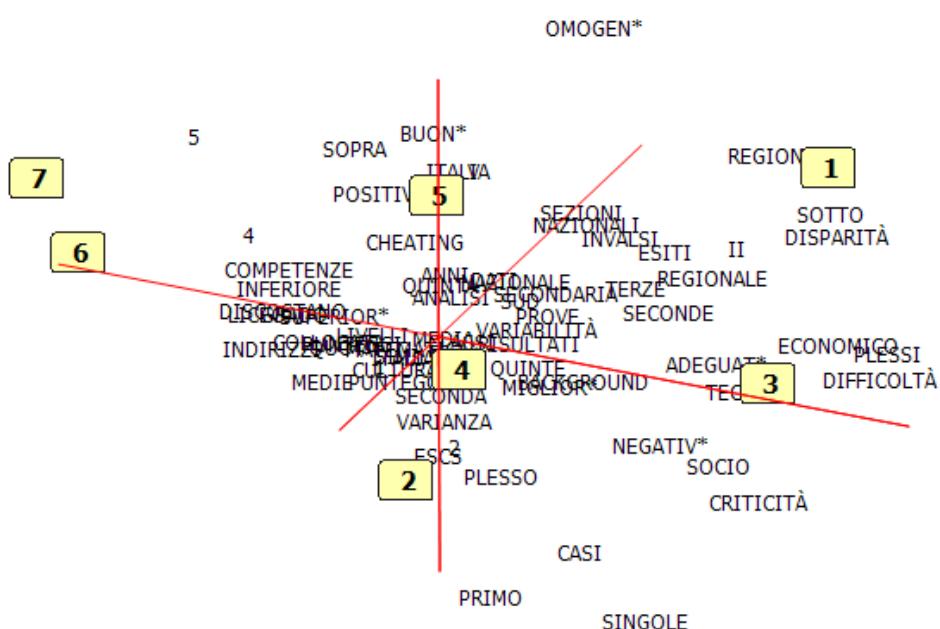
Nel terzo asse, infine, sul versante positivo si collocano gli istituti scolastici che si sono attribuiti il livello più basso di qualità, ossia il livello 1. Nelle motivazioni di queste scuole sono presenti diversi aggettivi che fanno pensare al confronto territoriale ("regionale", "regionali", "nazionale", "nazionali"). Risulta rilevante anche il contributo dei termini "sotto" e "tecnico". Nel semiasse negativo si

concentrano invece alcune parole che non risultano associate a specifici sottogruppi di scuole (i termini “competenze”, “varianza”, “indirizzi”, “grado”, “primo”, “plesso”, “sud”, “medie”)?<sup>7</sup>

Il confronto tra sottogruppi di scuole fa emergere un quadro sostanzialmente coerente con le finalità del percorso di autovalutazione, evidenziato sia dalla differente distribuzione dei termini a valenza positiva e negativa nelle motivazioni delle scuole collocate sui due versanti estremi della scala di autovalutazione, sia dalla presenza di

parole riprese dai descrittori e coerenti con il livello di qualità autoattribuito. L’analisi delle corrispondenze, inoltre, conferma la trasversalità rispetto ai gruppi delle parole associate alle categorie relative all’analisi dei dati, al confronto interno alla scuola e al confronto territoriale. Infine, ha permesso di comprendere i contesti di utilizzo di parole come l’aggettivo “inferiore”, che seppur presentando un significato negativo viene utilizzato prevalentemente per esprimere un senso di positività nelle motivazioni delle scuole collocate nei livelli più elevati della scala di valutazione.

**Figura 5.3 - Analisi delle corrispondenze lessicali (su parole del vocabolario tematico con  $p < 0,050$ )**



<sup>7</sup> Per rendere maggiormente leggibile la fig. 7.1 elenchiamo di seguito le parole che si affollano vicino all’origine degli assi e quelle che risultano nascoste. Le parole che si collocano vicino all’origine degli assi sono: classi, collocata, culturale, dati, discostano, italiano,

livelli, matematica, media, miglior\*, prove, punteggi, punteggio, quota, seconda, secondaria, simile, superior\*, variabilità. Le parole non visibili nella figura sono: 1, analisi, anni, background, grado, liceo, medie, nazionale, singoli, sud.

## RISULTATI IN SINTESI

- L'analisi testuale delle Motivazioni del giudizio dell'area Risultati nelle prove standardizzate nazionali ha fornito elementi per rispondere alle domande di ricerca che hanno orientato il progetto di metavalutazione del RAV. Rispetto all'efficacia degli strumenti di autovalutazione nell'orientare le scuole verso l'utilizzo di un approccio all'autovalutazione basato sui dati, i risultati dell'analisi evidenziano che circa la metà delle scuole ha fatto esplicito riferimento nelle motivazioni alla discussione o presentazione dei risultati per giustificare il livello di qualità autoattribuito. Al di là della diffusione delle forme grafiche di tipo numerico, l'utilizzo di parole che rimandano a un'analisi interna alla scuola o al confronto con valori di riferimento territoriale lasciano presupporre una riflessione effettuata a partire dai dati disponibili.
- Inoltre, dall'analisi testuale emergono informazioni che supportano l'idea che le scuole abbiano espresso giudizi rispetto a tutti i temi indicati come rilevanti dal Quadro di riferimento. Il tema trattato con minore frequenza è la varianza tra le classi, che tuttavia compare nel 60% delle motivazioni. Questo risultato potrebbe dipendere da una minore familiarità delle scuole con il concetto di varianza e con la sua interpretazione, come è emerso in altre analisi (cfr. cap. 3).
- Il segmento di percorso di autovalutazione analizzato, risulta coerente in quanto, come evidenziato dalla differente frequenza dei termini a valenza positiva e negativa, i testi delle motivazioni risultano coerenti con la scelta dei livelli di qualità nella scala di valutazione.
- Dall'analisi non emergono temi innovativi rispetto a quelli introdotti dagli strumenti di valutazione. Emerge, piuttosto, un approccio autonomo all'analisi dei dati, proposto dalle scuole, che valorizza il confronto territoriale subnazionale, utilizzato da più del 30% delle scuole. Ciò evidenzia l'esigenza di una più specifica contestualizzazione dell'analisi relativa ai risultati delle prove nazionali che attualmente non è prevista dallo strumento di autovalutazione.

## RIFLESSIONI E PROSPETTIVE

- In generale i risultati confermano un quadro di sostanziale validità del processo di autovalutazione poiché le scuole hanno riflettuto sugli aspetti che compongono il costrutto di qualità definito teoricamente.
- Nel valutare l'impatto dello strumento, e in particolare delle rubriche di valutazione, occorre però considerare anche l'influenza che i termini e le espressioni utilizzate nella rubrica esercitano sulle motivazioni delle scuole. Nell'area in esame, il vocabolario delle rubriche ha rappresentato prevalentemente un repertorio dal quale le scuole hanno attinto termini ed espressioni di carattere specialistico. Questo risultato può essere considerato come in parte atteso in una situazione -, come quella attuale -, nella quale, con l'avvio del Sistema nazionale di valutazione, si va costituendo anche un nuovo lessico per le scuole in campo valutativo...

## Appendice delle tabelle

**Tabelle 5.8 - Parole a frequenza elevata presenti nella sezione Motivazione del giudizio dell'area Risultati nelle prove standardizzate nazionali**

	FREQ.	% FREQ.	N. CASI	% CASI	TF • IDF
E	3249	4,9%	609	84,0%	246
DI	2897	4,3%	630	86,9%	176,7
IN	2739	4,1%	590	81,4%	245,1
MATEMATICA	1480	2,2%	505	69,7%	232,4
LA	1464	2,2%	570	78,6%	152,9
ITALIANO	1450	2,2%	504	69,5%	229
CLASSI	1427	2,1%	532	73,4%	191,8
MEDIA	1301	1,9%	484	66,8%	228,3
SCUOLA	1052	1,6%	510	70,3%	160,7
È	1037	1,6%	428	59,0%	237,4
DELLA	958	1,4%	489	67,5%	163,8
A	919	1,4%	471	65,0%	172,1
CON	908	1,4%	501	69,1%	145,7
SI	873	1,3%	467	64,4%	166,8
I	815	1,2%	495	68,3%	135,1
IL	807	1,2%	514	70,9%	120,5
NAZIONALE	718	1,1%	434	59,9%	160
DELLE	712	1,1%	461	63,6%	140
ALLA	676	1,0%	397	54,8%	176,8
PROVE	665	1,0%	514	70,9%	99,3
CHE	640	1,0%	351	48,4%	201,6
LE	607	0,9%	307	42,3%	226,5
SUPERIORE	604	0,9%	355	49,0%	187,3
PER	573	0,9%	266	36,7%	249,5
STUDENTI	566	0,8%	411	56,7%	139,5
TRA	544	0,8%	428	59,0%	124,5
LIVELLI	531	0,8%	404	55,7%	134,9
INFERIORE	515	0,8%	333	45,9%	174
INVALSI	504	0,8%	434	59,9%	112,3
NON	486	0,7%	354	48,8%	151,3
SONO	486	0,7%	325	44,8%	169,3

FONTE: elaborazione Invalsi, dati RAV a.s. 2014/2015

**Tabella 5.9 - Parole a frequenza media presenti nella sezione Motivazione del giudizio dell'area Risultati nelle prove standardizzate nazionali**

	FREQ.	% FREQ.	N. CASI	% CASI	TF • IDF
2	475	0,7%	370	51,0%	138,8
RISULTATI	463	0,7%	287	39,6%	186,3
1	455	0,7%	372	51,3%	131,9

<i>SCUOLE</i>	441	0,7%	405	55,9%	111,5
<i>ALLE</i>	438	0,7%	376	51,9%	124,9
<i>PUNTEGGIO</i>	425	0,6%	378	52,1%	120,2
<i>NEI</i>	421	0,6%	362	49,9%	127
<i>BACKGROUND</i>	410	0,6%	380	52,4%	115
<i>VARIANZA</i>	404	0,6%	359	49,5%	123,3
<i>SOCIO</i>	395	0,6%	375	51,7%	113,1
<i>SIMILE</i>	393	0,6%	372	51,3%	113,9
<i>CULTURALE</i>	385	0,6%	367	50,6%	113,8
<i>DISCOSTANO</i>	381	0,6%	252	34,8%	174,9
<i>ECONOMICO</i>	378	0,6%	362	49,9%	114
<i>QUELLO</i>	371	0,6%	339	46,8%	122,5
<i>QUOTA</i>	369	0,6%	324	44,7%	129,1
<i>PUNTEGGI</i>	365	0,5%	314	43,3%	132,6
<i>LINEA</i>	349	0,5%	261	36,0%	154,9
<i>UNA</i>	340	0,5%	202	27,9%	188,7
<i>LIVELLO</i>	340	0,5%	194	26,8%	194,7
<i>UN</i>	338	0,5%	207	28,6%	184
<i>COLLOCATA</i>	334	0,5%	301	41,5%	127,5
<i>RISPETTO</i>	328	0,5%	209	28,8%	177,2
<i>DALLA</i>	321	0,5%	291	40,1%	127,3
<i>QUELLA</i>	314	0,5%	285	39,3%	127,3
<i>DEL</i>	314	0,5%	171	23,6%	197
<i>NELLE</i>	302	0,5%	206	28,4%	165
<i>DEI</i>	286	0,4%	204	28,1%	157,5
<i>NELLA</i>	285	0,4%	172	23,7%	178,1
<i>AL</i>	265	0,4%	157	21,7%	176,1
<i>PRIMARIA</i>	263	0,4%	168	23,2%	167
<i>DEGLI</i>	260	0,4%	177	24,4%	159,2
<i>O</i>	250	0,4%	192	26,5%	144,3
<i>L</i>	237	0,4%	167	23,0%	151,1
<i>NEL</i>	234	0,4%	162	22,3%	152,3
<i>ALUNNI</i>	229	0,3%	146	20,1%	159,4
<i>PIÙ</i>	225	0,3%	141	19,5%	160
<i>ANCHE</i>	219	0,3%	183	25,2%	130,9
<i>DATI</i>	214	0,3%	149	20,6%	147,1
<i>MOLTO</i>	209	0,3%	179	24,7%	127
<i>DELL</i>	196	0,3%	146	20,1%	136,4
<i>SIA</i>	195	0,3%	120	16,6%	152,3
<i>ISTITUTO</i>	194	0,3%	146	20,1%	135
<i>GLI</i>	187	0,3%	144	19,9%	131,3
<i>SECONDARIA</i>	184	0,3%	135	18,6%	134,3
<i>POCO</i>	169	0,3%	157	21,7%	112,3
<i>ESITI</i>	150	0,2%	116	16,0%	119,4
5	149	0,2%	104	14,3%	125,7

<i>SE</i>	144	0,2%	135	18,6%	105,1
<i>DA</i>	142	0,2%	111	15,3%	115,7
<i>NAZIONALI</i>	140	0,2%	110	15,2%	114,7
<i>ALCUNE</i>	137	0,2%	128	17,7%	103,2
<i>CASI</i>	134	0,2%	122	16,8%	103,7
<i>POSITIVO</i>	134	0,2%	121	16,7%	104,2
<i>CLASSE</i>	130	0,2%	93	12,8%	115,9
<i>MENTRE</i>	125	0,2%	102	14,1%	106,5
<i>NEGATIVO</i>	121	0,2%	119	16,4%	95
<i>DECISAMENTE</i>	119	0,2%	106	14,6%	99,4
<i>QUANTO</i>	118	0,2%	95	13,1%	104,1
<i>SECONDE</i>	115	0,2%	97	13,4%	100,5
<i>4</i>	113	0,2%	87	12,0%	104,1
<i>ALL</i>	111	0,2%	90	12,4%	100,6
<i>SUPERIORI</i>	111	0,2%	86	11,9%	102,8
<i>RISULTA</i>	109	0,2%	78	10,8%	105,5
<i>CI</i>	104	0,2%	103	14,2%	88,1
<i>QUINTE</i>	104	0,2%	82	11,3%	98,4
<i>INFERIORI</i>	104	0,2%	79	10,9%	100,1
<i>AI</i>	104	0,2%	76	10,5%	101,9
<i>ED</i>	102	0,2%	94	13,0%	90,5
<i>QUELLI</i>	101	0,2%	82	11,3%	95,6
<i>DATO</i>	101	0,2%	72	9,9%	101,3
<i>SINGOLE</i>	100	0,2%	98	13,5%	86,9
<i>AD</i>	100	0,2%	80	11,0%	95,7
<i>GRADO</i>	98	0,2%	81	11,2%	93,3
<i>STANDARDIZZATE</i>	97	0,1%	89	12,3%	88,4
<i>REGIONALE</i>	97	0,1%	74	10,2%	96,1
<i>PERCENTUALE</i>	96	0,1%	67	9,2%	99,3
<i>DIVERSE</i>	95	0,1%	90	12,4%	86,1
<i>CRITICITÀ</i>	95	0,1%	79	10,9%	91,5
<i>3</i>	95	0,1%	58	8,0%	104,2
<i>MA</i>	91	0,1%	82	11,3%	86,1
<i>MEDIE</i>	88	0,1%	62	8,6%	94
<i>COMPETENZE</i>	84	0,1%	66	9,1%	87,4
<i>VALUTAZIONE</i>	84	0,1%	64	8,8%	88,5
<i>RIFERIMENTO</i>	81	0,1%	58	8,0%	88,8
<i>PROVA</i>	81	0,1%	55	7,6%	90,7
<i>COME</i>	81	0,1%	50	6,9%	94,1
<i>HA</i>	77	0,1%	61	8,4%	82,8
<i>MEDIO</i>	76	0,1%	55	7,6%	85,1
<i>VARIABILITÀ</i>	76	0,1%	54	7,5%	85,7
<i>ANALISI</i>	74	0,1%	65	9,0%	77,5
<i>SITUAZIONE</i>	74	0,1%	64	8,8%	78
<i>APPRENDIMENTO</i>	73	0,1%	53	7,3%	82,9

<i>DUE</i>	72	0,1%	59	8,1%	78,4
<i>NELL</i>	68	0,1%	57	7,9%	75,1
<i>ITALIA</i>	66	0,1%	39	5,4%	83,8
<i>LICEO</i>	66	0,1%	36	5,0%	86,1
<i>PRIMO</i>	65	0,1%	60	8,3%	70,3
<i>ALCUNI</i>	64	0,1%	60	8,3%	69,3
<i>TUTTI</i>	62	0,1%	58	8,0%	68
<i>GIUDIZIO</i>	61	0,1%	60	8,3%	66
<i>UGUALE</i>	61	0,1%	60	8,3%	66
<i>SOLO</i>	60	0,1%	54	7,5%	67,7
<i>HANNO</i>	60	0,1%	48	6,6%	70,7
<i>PARTE</i>	59	0,1%	54	7,5%	66,5
<i>CHEATING</i>	59	0,1%	45	6,2%	71,2
<i>SOPRATTUTTO</i>	58	0,1%	50	6,9%	67,4
<i>SUD</i>	58	0,1%	32	4,4%	78,6

FONTE: elaborazione Invalsi, dati RAV a.s. 2014/2015