

Riflettere sulle parole: la formazione delle parole nelle prove Invalsi

Zuzana Tóth

ABSTRACT

The present study examines the word formation tasks administered in the language awareness section of the Invalsi tests, starting from 2008, the first year when the tests were administered, up to 2019. The objective of this revision is to gain a deeper understanding of 1) which word formation processes (e.g., derivation, composition, conversion, etc.) are put under scrutiny in the tasks; 2) how the tasks are distributed between school levels and levels of difficulty; 3) which variables affect task difficulty; 4) if the test results allow for the identification of a progression of topics, from the easiest to the most difficult. The analyses reveal that most of the tasks investigate the process of derivation and that their difficulty is influenced by the interaction of several variables. The most salient variables are the complexity of the task (number of words to analyse and the format), the degree of abstraction required by the question and the frequency of the words to analyse.

1. INTRODUZIONE

Come osservano numerosi studi (es. Iacobini 2010; Grossmann – Rainer 2004; Lo Duca 2020), la formazione delle parole è un insieme di procedimenti che permettono di generare parole nuove e dunque di arricchire il lessico di una lingua a partire dal materiale linguistico preesistente. Le parole già presenti in una lingua possono essere modificate e acquistare significati nuovi attraverso procedimenti come l'aggiunta di affissi (*dis+onesto* → *disonesto*, *letam[e]+ aio* → *letamaio*), l'unione di più parole (*pesce+cane* → *pescecane*), il cambiamento della categoria lessicale (*mangiare* [verbo] → *mangiata* [no-

me]), e così via. I diversi procedimenti di formazione delle parole possono essere classificati in cinque categorie principali in base al tipo di cambiamento formale che avviene nel passaggio dalla parola di base alla parola nuova: prefissazione, suffissazione, parasintesi, composizione e conversione.

Come evidenzia Lo Duca (2020: 13), la formazione delle parole è un tema interessante sia da una prospettiva sincronica sia da una prospettiva diacronica, dato che l'italiano ha ereditato dal latino classico o tardo non solo la maggior parte del lessico, ma anche «procedimenti derivativi e affissi, i quali sono spesso penetrati in tutto il territorio romanzo, subendo vari tipi di adattamenti alle proprietà fono-morfologiche delle diverse lingue». Oltre ai procedimenti derivativi, l'italiano ha ereditato anche molte parole derivate e le rispettive parole di base. Per questo motivo, come spiega Lo Duca (2020: 13), al parlante di oggi potrebbe sembrare che alcune parole derivate che l'italiano ha ereditato dal latino (come *bibliotecario*, da lat. *bibliothecarius*), si siano formate in italiano, dato che è facile scomporle in una parola di base (*biblioteca*, dal lat. *bibliotheca*) e un suffisso (-ario dal lat. -arius).

La descrizione di questi procedimenti complessi non fa parte degli obiettivi del presente lavoro. Per approfondire questo tema, rimandiamo alla vasta letteratura esistente (es. Grossmann – Rainer 2004; Lo Duca 2020; Micheli 2020).

Il presente studio si occupa dei quesiti di formazione delle parole nelle prove Invalsi, quesiti che indagano in primo luogo il fenomeno della derivazione, in particolare la prefissazione e la suffissazione, e mettono alla prova la capacità degli studenti di riconoscere i procedimenti di formazione delle parole e i loro esiti semantici. Si propone di analizzare i quesiti somministrati negli undici anni in cui le prove si sono svolte regolarmente (dal 2008 al 2019), per capire meglio quali siano i temi più facili e più difficili per gli studenti.

2. LA FORMAZIONE DELLE PAROLE NELLE PROVE INVALSI

I compiti che riguardano la formazione delle parole sono inseriti in una sezione apposita delle prove Invalsi, dedicata alla riflessione sulla lingua. Come spiegato nel Quadro di Riferimento delle prove di Italiano (QdR)¹, i quesiti sono formulati con l'obiettivo di sollecitare l'osservazione e l'analisi di dati linguistici (parole, frasi o brevi testi) e di incoraggiare l'interrogazione della propria sensibilità linguistica. Per questo motivo nella formulazione dei quesiti di formazione delle parole si assume una prospettiva sincronica, «che guarda all'oggi con la sensibilità del parlante di oggi» (Lo Duca 2020: 14). Ad esempio, le prove non contengono quesiti volti a testare se gli alunni riescono a distinguere le parole che si sono formate in italiano (es. *bancario*, *universitario*) da quelle ereditate in forma derivata dal latino (es. *bibliotecario*, *segretario*), dato che entrambi i gruppi di parole presentano una base e un morfema derivativo chiaramente riconoscibili in italiano. Questo tipo di distinzione non sarebbe infatti ricostruibile in base alla naturale sensibilità linguistica del parlante di oggi.

Secondo il QdR (p. 12), i quesiti sulla formazione delle parole vertono su temi come «parole semplici e parole complesse; parole di base e parole derivate; parole alterate; parole composte; parole polirematiche». Non è però esplicitato il modo in cui questi temi sono scanditi nei diversi gradi scolastici. Infatti, il documento spiega che «le disposizioni ministeriale vigenti non consentono allo stato attuale di ricavare un sillabo dettagliato ed esplicito in materia grammaticale» (ivi, p. 13).

Per convincerci della genericità delle indicazioni ministeriali basta leggere le *Indicazioni nazionali per il primo ciclo d'istruzione*², nelle quali gli obiettivi di apprendimento in materia di formazione delle parole al termine della V primaria, cioè la conoscenza dei «principali mec-

¹ Documento disponibile in rete al seguente indirizzo: https://invalsi-areaprove.cineca.it/docs/file/qdr_italiano.pdf.

² Documento disponibile in rete al seguente indirizzo: http://www.indicazioni nazionali.it/wp-content/uploads/2018/08/Indicazioni_Annali_Definitivo.pdf.

canismi di formazione delle parole (parole semplici, derivate e composte)» vengono riproposti con una formulazione solo leggermente diversa al termine della III secondaria di I grado: «conoscere i principali meccanismi di formazione delle parole: derivazione, composizione».

Dunque, per disporre i contenuti nei diversi livelli scolastici e lungo una scala di progressiva difficoltà, l'Invalsi deve integrare le indicazioni ministeriali con informazioni provenienti da altre fonti, come la ricerca sulla consapevolezza linguistica, l'esperienza degli autori delle prove e degli esperti disciplinari, nonché i risultati dei rilevamenti progressivi (QdR, p. 13).

Nel caso della formazione delle parole però i suggerimenti della ricerca sono pochi. Differentemente da altri ambiti della lingua italiana, per i quali sono disponibili numerosi studi che possono essere interpretati in chiave didattica (si veda la vasta letteratura psicolinguistica sull'interpretazione delle espressioni anaforiche [es. Brown-Schmidt *et al.* 2005; Cacciari *et al.* 1997; Fukumara – van Gompel 2015; Järvikivi *et al.* 2005; Shake – Stine Morrow 2012]), la formazione delle parole è un ambito esplorato per lo più da un punto di vista teorico e descrittivo, con poche proposte su come le conoscenze teoriche possano essere integrate nell'insegnamento (Lo Duca 2021).

Il presente articolo si pone l'obiettivo di indagare in che modo gli studenti di diversi livelli scolastici hanno risposto ai quesiti Invalsi sulla formazione delle parole, nel periodo tra il 2008, l'anno in cui per la prima volta ha avuto luogo la rilevazione nazionale in III secondaria di I grado, e il 2019, l'ultimo anno precedente la pandemia, durante il quale le prove sono state somministrate regolarmente. I quesiti saranno analizzati con l'obiettivo di rispondere alle seguenti domande:

1. Come si distribuiscono i quesiti lungo una scala di difficoltà nei diversi gradi scolastici?
2. Quali sono le caratteristiche comuni dei quesiti che si collocano sullo stesso livello di difficoltà?

Per rispondere alla seconda domanda è necessario integrare i risultati delle prove Invalsi con un'analisi qualitativa dei quesiti. Il metodo di analisi adoperato nel presente studio è stato ideato secondo i principi dell'analisi qualitativa del contenuto (Mayring 2014). L'analisi dei quesiti è stata guidata da una griglia predefinita, composta da tre categorie principali articolate in dieci sottocategorie (Tabella 1). Tale categorizzazione è stata guidata dalle nostre domande di ricerca e da alcuni studi precedenti sulle prove Invalsi (es. Lo Duca – Toth 2021; Toth 2016).

Macrocategorie	Sottocategorie
1. Formato del quesito	1a. Scelta multipla semplice
	1b. Scelta multipla complessa
	1c. Risposta aperta univoca
2. Tipo di procedimento indagato	2a. Derivazione
	2b. Alterazione
	2c. Composizione
	2d. Parole polirematiche
3. Frequenza delle parole nel linguaggio	3a. Vocabolario di base
	3b. Lessico comune
	3c. Lessico tecnico-specialistico

Tabella 1. Sistema di categorie per l'analisi dei quesiti.

La categoria del formato si articola in tre sottocategorie, definite in base al QdR (p. 14-15). I tipi di procedimento di formazione delle parole sono state selezionate in base a un'analisi preliminare dei quesiti, con l'ausilio della banca dati elettronica dell'Invalsi³. La frequenza delle parole nel linguaggio è stata stabilita in base all'attuale versione online del *Vocabolario di base della lingua italiana*⁴. Si tratta di un vocabolario redatto da Tullio De Mauro nel 1980 e da allora più volte rivisto e aggiornato, che raccoglie circa 7000 parole, suddivise in tre categorie in base alla loro frequenza d'uso in un campione rappresentativo di testi

³ Consultabile all'indirizzo: www.gestinv.it.

⁴ Consultabile all'indirizzo: <https://www.internazionale.it/opinione/tullio-de-mauro/2016/12/23/il-nuovo-vocabolario-di-base-della-lingua-italiana>.

scritti e registrazioni del parlato. Nella versione più aggiornata, presentata da De Mauro (2016), la categoria del *vocabolario fondamentale* comprende circa 2000 parole di altissima frequenza, le cui occorrenze costituiscono circa l'86% delle occorrenze lessicali nei testi sottoposti all'analisi. La categoria del *vocabolario di alto uso* raccoglie circa 3000 parole di uso nettamente minore rispetto al vocabolario fondamentale, le cui occorrenze costituiscono circa il 6% per cento delle occorrenze lessicali nei testi analizzati. Tuttavia, queste parole sono di uso significativamente maggiore rispetto alle parole del lessico comune e alla restante «sterminata massa del lessico» (De Mauro 2016). La categoria del *vocabolario di alta disponibilità* comprende circa 2500 parole che, nonostante la loro scarsa frequenza nel parlato e nello scritto, sono ben note ai parlanti perché legate «ad atti e oggetti di grande rilevanza nella vita quotidiana (*alluce, batuffolo, carrozzeria, dogana, ecc.*)» (De Mauro 2004: XI). Alle parole del vocabolario di base si aggiungono quelle del lessico comune, a cui appartengono «vocaboli che sono usati e compresi indipendentemente dalla professione o mestiere che si esercita o dalla collocazione regionale e che sono generalmente noti a chiunque abbia un livello mediosuperiore di istruzione» (De Mauro 2004: XI) e quelle del lessico tecnico-specialistico, che sono «legati a un uso marcatamente o esclusivamente tecnico o scientifico e noti soprattutto in rapporto a particolari attività, tecnologie, scienze» (ivi).

3. ANALISI DEI DATI

La Tabella 2 offre una panoramica dei quesiti di formazione delle parole somministrati nei diversi livelli scolastici. Come si vede, dal 2008 al 2019 sono stati somministrati 47 quesiti, di cui la maggioranza (36) ha riguardato il procedimento della derivazione. Sono meno numerosi i quesiti inerenti a temi come l'alterazione (2), la composizione (6) o le parole polirematiche (3).

La distribuzione dei quesiti nei diversi gradi scolastici⁵ è abbastanza equilibrata. Il numero limitato dei quesiti di I secondaria di I grado (G6) è dovuto al fatto che al G6 la somministrazione è avvenuta solo quattro volte, dall'a.s. 2009/2010 all'a.s. 2012/2013. Dato il numero molto ridotto dei quesiti somministrati in questo grado scolare, vi dedicheremo solo un'attenzione marginale. Le leggere differenze nel numero dei quesiti tra i restanti tre gradi scolastici sono attribuibili al fatto che nel primo anno di somministrazione (2008) le prove hanno coinvolto solo il grado 8, con l'aggiunta del grado 5 e 10 rispettivamente nel 2009 e nel 2012.

	G5	G6	G8	G10	Totale
Derivazione	10	4	13	9	36
Alterazione	2	0	0	0	2
Composizione	2	0	2	2	6
Polirematiche	0	0	2	1	3
Totale	14	4	17	12	

Tabella 2. Panoramica dei quesiti di formazione delle parole nelle prove Invalsi.

Nonostante la ricchezza dei dati quantitativi raccolti dall'Invalsi, la disposizione dei quesiti lungo una scala di difficoltà è stato uno degli aspetti problematici di questa ricerca, per due motivi.

Il primo è collegato alla modalità di somministrazione delle prove e all'elaborazione statistica dei dati. Come evidenziato anche in Lo Duca e Toth (2021: 19), fino all'anno scolastico 2016/2017 tutte le prove sono state somministrate in formato cartaceo e rese pubbliche al termine della rilevazione. Dall'anno scolastico 2017/2018 invece le prove della scuola secondaria di primo e secondo grado vengono somministrate al computer e restano segrete, mentre sono rese pubbliche quelle della scuola primaria, somministrate tuttora su carta. Le differenze nella modalità di somministrazione comportano anche delle differenze nel-

⁵ L'indicazione del grado scolastico avviene attraverso l'indicazione dei rispettivi anni di scolarizzazione. Grado 5 indica la V primaria, grado 8 la III secondaria di I grado, ecc.

l'elaborazione statistica dei dati (per una discussione più dettagliata di questo punto rimandiamo a Desimoni 2018; Lo Duca – Toth 2021: 19). Per questo motivo, nonostante la disponibilità di metodi statistici più complessi e precisi, in questo lavoro usiamo la percentuale di risposte corrette come bussola nell'identificazione del grado di difficoltà dei quesiti. Il motivo di questa scelta è semplice: la distribuzione percentuale delle risposte è l'unico dato disponibile per tutti i quesiti.

Il secondo problema riguarda la definizione della risposta corretta nei quesiti che hanno un formato complesso, ad esempio contengono tabelle da riempire con molti item da analizzare. Talvolta, la risposta a questi quesiti viene considerata corretta anche se lo studente commette un errore, cioè dà una risposta errata a uno degli *item*. In questi casi dobbiamo chiederci se il quesito non sia in realtà più difficile di quello che suggerisce la percentuale di risposte corrette, soprattutto se è presente un item particolarmente difficile, che molti studenti avrebbero potuto sbagliare, senza che la percentuale di risposte corrette rispecchi questa difficoltà. Per questo motivo, anche la percentuale di risposte corrette può essere considerata solo un dato orientativo, che deve essere integrato con valutazioni qualitative nel caso di quesiti con un formato complesso.

3.1. La costruzione di una scala di difficoltà

I quesiti sono stati suddivisi in cinque gruppi di difficoltà, seguendo i criteri adottati in un lavoro precedente sull'ortografia nelle prove Invalsi (si veda Lo Duca – Toth 2021), riportati nella Tabella 3.

Livello di difficoltà	Percentuale di risposte corrette
Facile	più del 70%
Medio-facile	69 %-60%
Medio	59 %-40%
Medio-difficile	39 %-30%
Difficile	meno del 30%

Tabella 3. Criteri per la classificazione dei quesiti in base alla loro difficoltà.

Nella Figura 1 sono riportati i dati relativi alla distribuzione dei quesiti di formazione delle parole e di ortografia nei cinque livelli di difficoltà individuati.

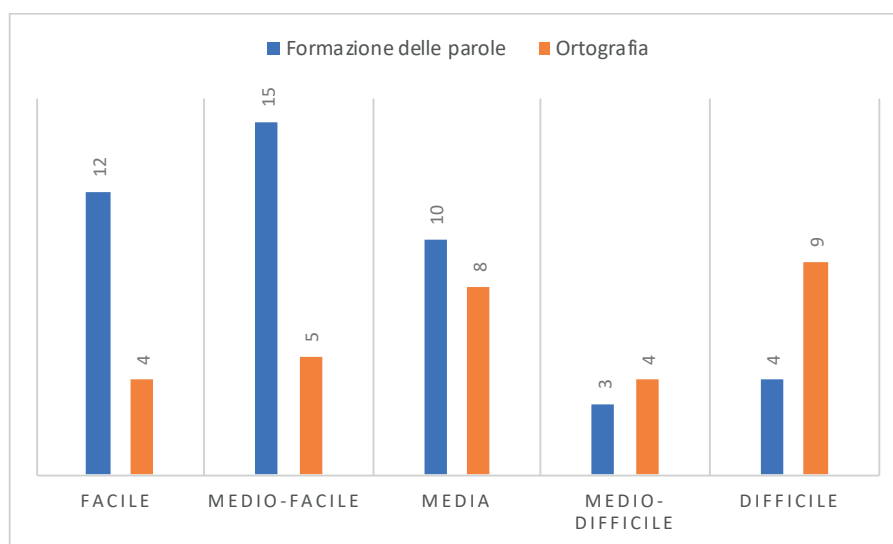


Figura 1. Distribuzione dei quesiti nei cinque livelli di difficoltà

Il primo risultato che emerge da questo confronto è la relativa facilità dei quesiti di formazione delle parole rispetto a quelli di ortografia. Come si evince dalla Figura 1., la maggior parte dei quesiti di formazione delle parole (il 61% del totale) è risultata facile (12) o medio-facile (15). Per quanto riguarda invece l'ortografia, solo il 30% (4+5) dei quesiti si colloca nella parte bassa della scala di difficoltà. La tendenza opposta è osservabile nella parte più alta della scala: solo 16% (cioè 7 quesiti su un totale di 44) dei quesiti di formazione delle parole risulta medio-difficile o difficile, mentre nel caso dell'ortografia la proporzione di quesiti medio-difficili e difficili è pari al 43% del totale.

In base a ricerche pregresse sui risultati delle prove Invalsi (es. Lo Duca 2018; Toth 2019, 2020; Lo Duca – Toth 2021) e sulla capacità di riflessione linguistica degli studenti (es. Lo Duca – Polato 2010; Lo Duca *et alii* 2011; Myhill 2000) è possibile ipotizzare che la relativa facilità dei quesiti di formazione delle parole sia dovuta a una serie di

motivi interconnessi. Il più importante sembra essere la focalizzazione dell'attenzione sul significato piuttosto che sulle caratteristiche formali o morfosintattiche delle parole. Come infatti mostrano gli studi sopra citati, la riflessione sul significato richiede un livello di astrazione minore e di conseguenza risulta quasi sempre più facile della riflessione sulla morfosintassi. Inoltre, come sarà illustrato anche su esempi concreti nel paragrafo 3, la riflessione sul significato permette di far affidamento sulla propria naturale competenza linguistica e non richiede necessariamente il ricorso a conoscenze linguistiche esplicite e verbalizzabili, trasmesse dall'insegnamento, come invece spesso avviene nei quesiti sulla morfosintassi o sull'ortografia. Anche il modo in cui sono formulati i quesiti agevola il ricorso alla competenza implicita e sollecita la riflessione sui dati linguistici, piuttosto che il ricorso a conoscenze pregresse. In genere, le conoscenze esplicite richieste dai compiti sono poche ed essenziali, per esempio la conoscenza di alcuni termini tecnici, come *suffisso*, *prefisso*, *derivazione*, *composizione*, *alterazione*, ecc.

3.2. La difficoltà dei quesiti e la frequenza delle parole nel linguaggio

Dato che i quesiti focalizzano l'attenzione sul significato di parole selezionate, sembra giusto chiedersi quale sia la relazione tra la difficoltà di un quesito e la frequenza delle parole su cui esso verte. Si potrebbe pensare che le parole del VdB, probabilmente più frequenti anche nell'esperienza linguistica degli studenti, siano più facili da analizzare di quelle del lessico comune, appena più impegnativo, o del lessico specialistico, ad eccezione, naturalmente, dei termini ricorrenti nelle discipline scolastiche.

I dati attualmente a disposizione non consentono di dare una risposta univoca a questa domanda.

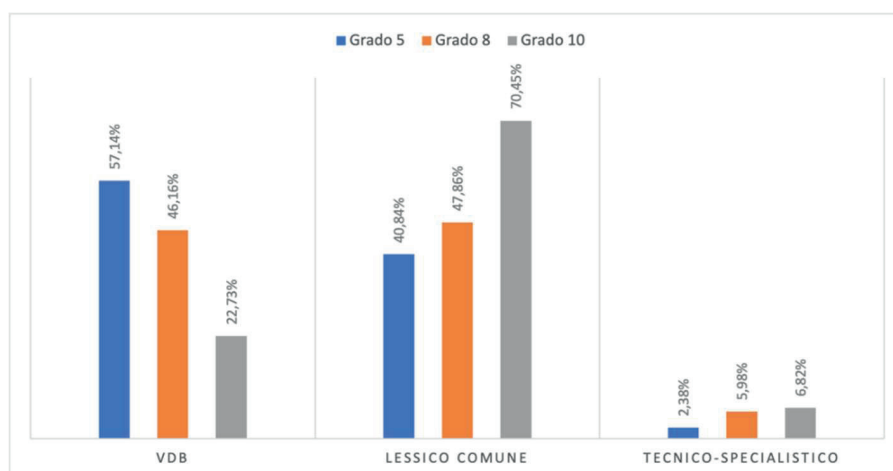


Figura 2. Distribuzione delle parole su cui vertono i quesiti.

La Figura 2 mostra la distribuzione percentuale delle 316 parole che finora sono state oggetto di analisi nei quesiti Invalsi accorpate in tre categorie, in base all'edizione online del dizionario *Il nuovo De Mauro*⁶: il lessico comune, il VdB e il lessico tecnico-specialistico. Una buona parte delle parole su cui vertono i quesiti non appartiene al VdB, ma al lessico comune, soprattutto nel grado 10 (70,45%). Nei gradi scolastici inferiori (G5 e G8) le parole si dividono in modo abbastanza equo tra il VdB (57,14% nel G5 e 46,16% nel G8) e il lessico comune (40,84% nel G5 e 47,86% nel G8). Solo una parte esigua delle parole rientra nel lessico tecnico-specialistico.

Analizziamo ora i dati scorporati per grado scolastico. Al grado 5 (Figura 3) i quesiti facili hanno coinvolto un totale di 28 parole, di cui 13 fanno parte del VdB, 15 del lessico comune. Nei quesiti medio-facili e in quelli di media difficoltà invece il numero di parole del VdB (rispettivamente 15 e 16) è superiore a quelle del lessico comune (rispettivamente 8 e 10).

⁶ Il dizionario è consultabile in rete: <https://dizionario.internazionale.it>.

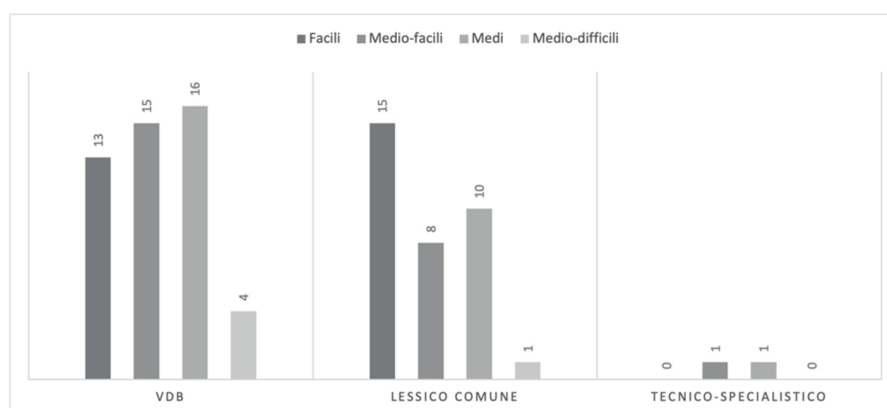


Figura 3. Distribuzione delle parole nei quesiti di G5⁷.

La distribuzione delle parole nei quesiti di grado 8 (Figura 4) è diversa. Nei quesiti facili e medio-facili le parole sottoposte all'analisi si distribuiscono in modo abbastanza equo tra il VdB (rispettivamente 26 e 24 parole) e il lessico comune (rispettivamente 21 e 17 parole), nei quesiti difficili invece prevalgono nettamente le parole del lessico comune (15) e ci sono solo 4 parole del VdB.

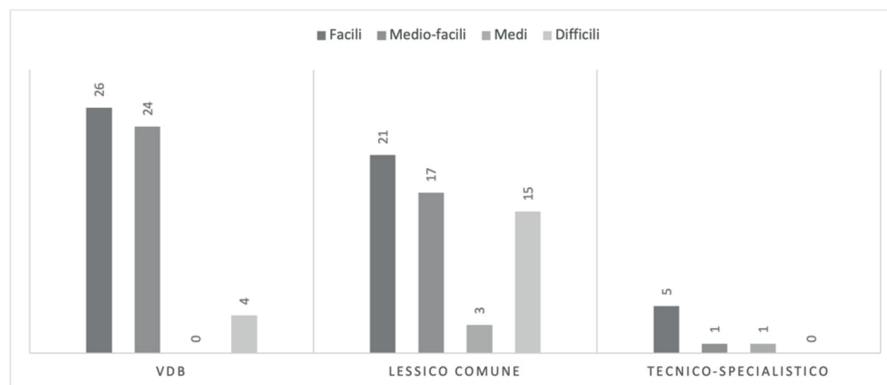


Figura 4. Distribuzione delle parole nei quesiti di G8⁸.

⁷ Nel nostro *corpus* non sono presenti quesiti difficili di G5, perciò tale categoria è assente nella figura 3.

⁸ Nel nostro *corpus* non sono presenti quesiti medio-difficili di G8, perciò tale categoria è assente nella figura 3.

Come già anticipato, nei quesiti di grado 10 la maggior parte delle parole appartiene al lessico comune, indipendentemente dalla difficoltà del quesito (si veda la Figura 2).

3.3. La difficoltà dei quesiti e il formato

Alcuni studi precedentemente condotti sui risultati delle prove Invalsi (es. Toth 2019; Lo Duca – Toth 2021) mostrano che i quesiti con un formato complesso, ad esempio quelli che contengono tabelle con molti item da analizzare, tendono a essere più difficili dei quesiti a scelta multipla semplice. Questa tendenza viene in parte confermata anche dai dati analizzati nel presente lavoro. Come si vede nella Figura 5, su un totale di 29 quesiti a scelta multipla semplice, 13 risultano facili, 10 medio-facili, che insieme formano quasi l'80% del totale. I quesiti a scelta multipla complessa sono di numero minore, in totale 14, ma coprono tutti i livelli di difficoltà, concentrandosi soprattutto nella parte medio-alta della scala.

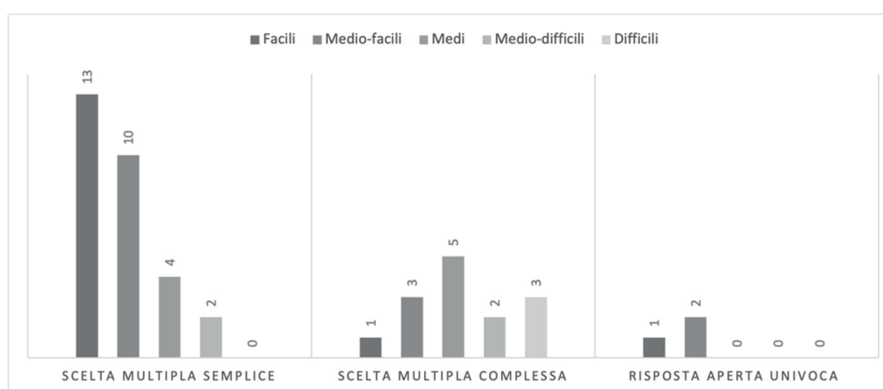


Figura 5. La relazione tra il formato dei quesiti e il loro livello di difficoltà.

Data l'esiguità dei dati a disposizione, non è possibile formulare generalizzazioni. Tuttavia, possiamo concludere che la frequenza delle parole nel linguaggio e il formato del quesito sono probabilmente due variabili che almeno in parte incidono sulla difficoltà dei quesiti. Nei

prossimi paragrafi useremo queste due variabili per descrivere, separatamente per ogni grado scolastico, i cinque livelli di difficoltà.

3.4. I quesiti di *V primaria*

La Tabella 4 offre uno sguardo d'insieme sui quesiti di formazione delle parole in *V primaria*. Le celle vuote rappresentano casi che per adesso non sono coperti da quesiti.

	Facili	Medio-facili	Medi	Medio-difficili	Difficili
Derivazione	C10512 C4519 C2518	C3515 C6511 C3516 C2510	C8516 C8514 C9511	C9517	
Alterazione	C5517				
Composizione	C2519 C7513				

Tabella 4⁹.

La maggior parte dei quesiti si colloca nella parte bassa della scala di difficoltà e riguarda il fenomeno della derivazione. La scarsa presenza di quesiti difficili probabilmente rispecchia la volontà di evitare contenuti troppo impegnativi a questo livello scolastico.

Una delle caratteristiche che accomunano i quesiti facili è la semplicità del formato e il numero ridotto di parole da analizzare. Ad esempio, nel caso del quesito C4519 è sufficiente capire che tutte e quattro le parole sono derivate, senza dover andare oltre e indicare, ad esempio, i diversi significati che i suffissi aggiungono alle parole di base.

⁹ Le sigle riportate nella tabella facilitano il ritrovamento dei quesiti nella banca dati pubblica dell'Invalsi. Il primo numero corrisponde al numero del quesito all'interno del fascicolo originario. Il secondo numero corrisponde al grado scolastico, mentre il terzo all'anno di somministrazione. Dunque, la sigla C10512 indica il quesito C10 nel fascicolo di *V primaria*, somministrato nel 2012. Per accedere a questa banca dati è necessario visitare il sito www.gestinv.it e registrarsi con un indirizzo e-mail.

- C4519. Che cos'hanno in comune questi quattro nomi?
ZUCCHERIERA – GIORNALISTA – FOGLIAME – GELATERIA
- A. Sono tutti nomi composti.
 - B. Sono tutti nomi collettivi.
 - C. Sono tutti nomi derivati.
 - D. Sono tutti nomi alterati.

Nei quattro quesiti medio-facili si richiede di osservare un maggior numero di parole (da sette a nove) e di identificarne le caratteristiche comuni. Ad esempio, due quesiti (C6511, C3515) richiedono di osservare parole appartenenti alla stessa famiglia lessicale (derivati di *acqua*, es. *acquazzone*, *acquario* e derivati di *carta*, es. *cartoccio*, *cartolina*) e di trovare l'intruso (nella prima serie *nacque* e nella seconda *cartilagine*). Il fatto che entrambi i quesiti ottengano percentuali abbastanza alte di risposte corrette (rispettivamente 66,9% e 68,4%) suggerisce che riconoscimento di caratteristiche comuni a una serie di parole e l'individuazione di un unico intruso sia un compito abbastanza facile. Questa ipotesi è confermata anche dagli altri quesiti di questo gruppo.

I quesiti di media difficoltà sono abbastanza diversi tra loro, perciò, con i dati attualmente a disposizione, è difficile proporre una descrizione di questo livello. Si tratta comunque di quesiti che richiedono di osservare un numero più elevato di parole (da cinque a sedici), per identificarne le caratteristiche comuni oppure per individuare un intruso.

Nell'unico quesito medio-difficile somministrato in V primaria (C9517) si richiede di individuare, per ciascuna delle cinque parole presenti in una tabella (*scommettere*, *sgonfiare*, *scoppiettare*, *scongolare*, *sprecare*), se la *s-* iniziale ha valore di prefisso privativo¹⁰ oppure no. Il quesito ha un formato complesso e richiede di analizzare singolarmente ogni *item*. Come verrà confermato più avanti, questo tipo di compito risulta spesso più difficile di quelli in cui è sufficiente mettere a fuoco le caratteristiche comuni a una serie di parole, senza classificarle singolarmente.

¹⁰ Il termine tecnico "prefisso privativo" viene spiegato nella consegna del quesito.

3.5. I quesiti di III secondaria di II grado

La Tabella 5 mostra che anche al grado 8 prevalgono i quesiti facili e medio-facili, che analizziamo in modo più dettagliato. L'unico quesito di difficoltà media, i due quesiti difficili e tutti gli altri indicati con una X e un numero progressivo non potranno essere commentati in modo dettagliato, perché fanno parte della banca dati secretata¹¹. Ci limiteremo quindi a dare indicazioni generiche relative al formato, al tipo di procedimento indagato o alla frequenza delle parole prese in esame.

	Facili	Medio-facili	Medi	Medio-difficili	Difficili
derivazione	D8812 C7811 X2 X3 X4	C1815 C8813 C7816 C10816 C2814 X5	X6		X7 X8
alterazione					
composizione	X1				
polirematiche	C8817	C4814			

Tabella 5.

Le caratteristiche più evidenti dei quesiti facili riguardanti la derivazione sono tre: la semplicità del formato, il numero limitato di parole

¹¹ A partire dall'anno scolastico 2017/2018 le prove della scuola secondaria di primo e secondo grado vengono somministrate al computer. Questa modalità di somministrazione comporta un cambiamento importante nella restituzione dei dati. Come si legge in un documento pubblicato dall'Invalsi (2022) sulle prove somministrate al computer, tali prove si compongono di domande estratte da un ampio repertorio di quesiti (*banca di item*) e variano da studente a studente, mantenendo per ciascun individuo uguale difficoltà e struttura. La banca di item viene progressivamente ampliata, ma una buona parte dei quesiti rimane invariata da anno in anno. Questo tipo di somministrazione consente di ottenere dati statisticamente più attendibili, ma rende impossibile la pubblicazione dei quesiti da anno in anno, al termine della rilevazione nazionale, come invece avviene per le prove cartacee, somministrate nella scuola primaria.

da analizzare e la semplicità del compito, che riguarda l'osservazione di caratteristiche salienti e facilmente identificabili. In tre casi lo studente deve analizzare quattro o cinque parole, per indicare l'unica parola derivata o l'unica parola composta. Due quesiti invece hanno un formato più complesso e vertono su un numero maggiore di parole (da 8 a 16). L'analisi però riguarda caratteristiche facilmente osservabili, ad esempio la distinzione tra parole derivate (es. *calcolatrice*, *legname*, *deprezzamento*, *urbanesimo*, *popolarità*) e parole di base (es. *disciplina*, *rischio*, *stipendio*).

I quesiti medio-facili di grado 8 richiedono di analizzare le parole in base a criteri più complessi, ad esempio in base al significato di un determinato prefisso o suffisso, come si vede nel quesito C3814, riportato sotto.

C3814. Indica in quali dei seguenti aggettivi *in-* è un prefisso con valore negativo e in quali non lo è.

Metti una crocetta per ogni riga.

Aggettivo	<i>in-</i> è un prefisso con valore negativo	<i>in-</i> non è un prefisso con valore negativo
a) indeciso		
b) intenso		
c) innamorato		
d) interno		
e) incapace		
f) integrale		
g) incivile		
h) insolubile		

Nonostante la complessità del compito, il quesito risulta medio-facile: il 60% degli studenti dà 8 risposte corrette su 8. La relativa facilità del quesito è probabilmente dovuta al fatto che tutte le parole tranne una (*indeciso*) appartengono al VdB. Il quesito X7, infatti, ha un'architettura simile, nel senso che richiede di analizzare il significato di uno stesso prefisso in 10 parole. In questo caso però solo una delle parole appartiene al VdB, tutte le altre fanno parte del lessico comune. Il tasso

di studenti che fornisce 10 risposte corrette su 10 al quesito X7 è molto inferiore, pari al 23%.

3.6. I quesiti di II secondaria di II grado

La Tabella 6 mostra che, diversamente dai gradi scolastici inferiori, i quesiti di grado 10 si concentrano nella parte centrale e alta della scala di difficoltà, forse perché gli autori delle prove hanno riservato i temi più impegnativi a questo livello scolastico.

	Facili	Medio-facili	Medi	Medio-difficili	Difficili
derivazione		E11015	E31011 E21013 E51015 X10 X11	X12 X13	E71017
alterazione					
composizione	E51014	X9			

Tabella 6.

Dato il ridotto numero di quesiti, non è possibile descrivere i primi due livelli.

I quesiti di media difficoltà condividono alcuni tratti dei quesiti difficili: hanno quasi tutti un formato complesso e richiedono di analizzare un numero abbastanza elevato di parole (da 6 a 13). Il quesito E31013, ad esempio, è composto da 3 item, ciascuno dei quali richiede di analizzare una parola prefissata (es. *prefestivo*) e di confrontarla con un gruppo di tre parole (es. *precotto*, *pregiato*, *premuroso*) per capire in quale di esse *pre-* ha lo stesso significato osservato nella prima parola. Un elemento che forse facilita l'individuazione della risposta corretta è che alcuni distrattori (come *pregiato* e *premuroso*) non contengono prefissi e sono dunque facili da escludere. Il quesito E51015 invece richiede di analizzare 6 parole e indicare ogni volta se il suffisso *-evole*

ha significato attivo (*favorevole*) o passivo (*maneggevole*), scegliendo tra parafrasi come *che favorisce* / *che può essere favorito*.

Il quesito E71017, l'unico difficile in questo grado scolastico, richiede di classificare otto parole terminanti in *-ino* in tre categorie, in base alla funzione del suffisso. Gli studenti devono cioè indicare se *-ino* forma 1) un diminutivo (come in *bastoncino*, *topolino*); 2) un nome di mestiere (es. *arrotino*, *spazzino*, *bagnino*); 3) un aggettivo (come in *mascolino*, *marino*, *settembrino*). Dunque, la difficoltà deriva probabilmente dalla compresenza di due variabili: la complessità del compito di isolare tre diversi significati di uno stesso suffisso e l'alto numero delle parole da analizzare. È inoltre necessario notare il grande squilibrio tra le risposte corrette ai singoli *item*: mentre la parola *topolino* (del VdB) registra il 93% di risposte corrette, *settembrino* (del lessico comune) è classificato come aggettivo solo dal 47%.

4. CONCLUSIONI

Dalle analisi presentate si evince che la difficoltà dei quesiti è influenzata da tre caratteristiche principali: la complessità formale del compito (cioè il numero di parole da analizzare e il formato del quesito), il grado di astrazione richiesto dal compito (cioè la focalizzazione dell'attenzione su caratteristiche facilmente osservabili delle parole, come la distinzione tra parole di base e parole derivate, o su caratteristiche più complesse, come i vari significati veicolati da uno stesso affisso) e la frequenza nel linguaggio delle parole proposte all'analisi. Queste variabili non hanno un effetto diretto sulla difficoltà del quesito: esse interagiscono tra di loro e la loro rilevanza cambia in funzione del grado scolastico. Un risultato abbastanza evidente è che i quesiti con un formato semplice, che richiedono di analizzare un numero limitato di parole, in base a criteri immediatamente riconoscibili, sono sempre più facili di quelli che contengono molti *item* da analizzare singolarmente, in base a criteri più astratti, più difficili da osservare. Tra questi due estremi si collocano i quesiti che contengono un numero maggiore di parole da analizzare, di cui però è necessario capire le caratteristiche comuni,

senza doverle classificare singolarmente. L'appartenenza delle parole al VdB oppure al lessico comune spiega soprattutto le differenze tra i quesiti nei diversi gradi scolastici e in alcuni casi la differenza nella difficoltà di quesiti strutturalmente molto simili.

Tuttavia, l'analisi dei dati a disposizione non consente di andare oltre. Per capire meglio quali sono i contenuti più facili e più difficili nei diversi livelli scolari sarebbe necessario allargare il nostro corpus di quesiti, per coprire in modo equilibrato non solo tutti i livelli di difficoltà, ma anche i diversi procedimenti (derivazione, alterazione, composizione ecc.) di formazione delle parole.

Università di Basilea
Ricercatrice Marie Curie
zuzana.toth@unibas.ch

BIBLIOGRAFIA

De Mauro, T.

2004 *Il Dizionario di Italiano Paravia*, Paravia, Torino.

2016 *Il nuovo vocabolario di base della lingua italiana*, in Id., *Guida all'uso delle parole*, Laterza, Roma – Bari, pp. 157-210.

Desimoni, M.

2018 *I livelli per la descrizione degli esiti delle prove INVALSI. Le rilevazioni degli apprendimenti (a.s. 2017-2018)*, online: https://invalsi-areaprove.cineca.it/docs/2018/Livelli_INVALSI_g8.pdf

Grossmann, M. – Rainer, F. (a cura di)

2004 *La formazione delle parole in italiano*, Max Niemeyer, Tübingen.

Iacobini, C.

2010 *Formazione delle parole*, in (a cura di), *Enciclopedia dell'italiano*, a cura di R. Simone, G. Berruto, P. D'Achille, Istituto dell'Enciclopedia italiana, Roma, vol. I, pp. 513-514.

Invalsi

2018 *Quadro di riferimento delle prove Invalsi di Italiano*, online: https://invalsi-areaprove.cineca.it/docs/file/qdr_italiano.pdf.

2022 *Le prove computer based (CBT). Ultimo anno secondaria di II grado (grado 13). A.S. 2022-2023*, online: https://invalsi-areaprove.cineca.it/docs/2023/Rilevazioni_Nazionali/2022-2023_Organizzazione_delle_prove_CBT_Grado_13.pdf

Lo Duca, M.G.

2020 *Italiano: la formazione delle parole*, Carocci, Roma.

2021 *Formazione delle parole e terminologia grammaticale: proposte per il mondo della scuola*, in «DIDIT», 1, pp. 11-32.

Lo Duca, M.G. – Cristinelli, A. – Martinelli, E.

2011 *Riconoscere le voci verbali: indagine su una categoria complessa*, in *Grammatica a scuola*, a cura di L. Corrà, W. Paschetto, Franco Angeli, Milano, pp. 153-171.

Lo Duca, M.G. – Ferronato, M. – Mengardo, M.

2009 *“Indicazioni per il Curricolo” e obiettivi di apprendimento sulle categorie lessicali: il riconoscimento del Nome*, in *Lingua e grammatica. Teoria e prospettive didattiche*, a cura di P. Baratter, S. Dallabrida, Milano, Franco Angeli, pp. 11-27.

Lo Duca, M.G. – Polato, S.

2010 *Dalle elementari alle soglie dell'università: indagine sul riconoscimento della categoria lessicale del nome*, in *Perché la*

grammatica? La didattica dell'italiano tra scuola e università, a cura di G. Fiorentino, Carocci, Roma, pp. 78-92.

Lo Duca, M.G. – Toth, Z.

2021 *La competenza ortografica nelle prove INVALSI*, in «Italiano a scuola», 3/1, pp. 1-38.

Micheli, M.S.

2020 *La formazione delle parole. Italiano e altre lingue*, Carocci, Roma.

Myhill, D.

2000 *Misconceptions and Difficulties in the Acquisition of Metalinguistic Knowledge*, in «Language and Education», 14/3, pp. 151-163.

Toth, Z.

2019 *La descrizione dei livelli di competenza in base alle prove INVALSI di Italiano*, in «Working Papers Invalsi», pp. 3-17.