

Il rapporto tra livello di difficoltà e livello di conoscenza della lingua nei testi delle prove di comprensione scritta della lingua italiana

Antonio Venturis

Aristotle University of Thessaloniki, Greece

Abstract This study aims to show the readability limits of the B and C language level Italian texts, which can be used for the reading assessment and the characteristics they depend on. For this purpose, 184 Italian texts used in the certification exams of the KPG system were analysed in an attempt to reveal the main traits (quantitative, lexical and syntactic) of the texts defining their readability and their correlation with the language level for which they are intended. The correlation analysis indicated that cognitive characteristics are more important than quantitative characteristics, while a high correlation of lexical variables with B1 and B2 levels and syntactic variables with C1 and C2 levels emerged.

Keywords Readability limits. Language levels. Written comprehension assessment. Language certification.

Sommario 1 Introduzione. – 2 Metodi di misurazione della leggibilità. – 3 Caratteristiche metodologiche della ricerca. – 4 Risultati della ricerca. – 5 Conclusioni.



Peer review

Submitted	2021-02-25
Accepted	2022-01-14
Published	2022-03-22

Open access

© 2022 | Creative Commons Attribution 4.0 International Public License



Citation Venturis, A. (2022). "Il rapporto tra livello di difficoltà e livello di conoscenza della lingua nei testi delle prove di comprensione scritta della lingua italiana". *EL.LE*, 11(1), 35-52.

DOI 10.30687/ELLE/2280-6792/2022/01/002

1 Introduzione

La grande importanza che la società greca attribuisce alla Certificazione di conoscenza della lingua straniera è ormai un dato di fatto, come è stato sottolineato da un buon numero di studiosi (Alexiou 2003, Dendrinos, Zouganelli, Karavas 2013; Psitos, Marotta 2018; Tsiplakidis 2014). Il fatto che gli allievi, sin dalle prime classi del ginnasio o ancora prima, siano orientati a ottenere la Certificazione è associato alla convinzione che essa costituisca un requisito fondamentale e necessario per un futuro sbocco professionale, ma è collegato anche a tutto ciò che è coinvolto nella ricerca del lavoro, come l'istruzione, l'emigrazione o l'acquisizione della cittadinanza (Hill, McNamara 2015). Questa convinzione appare più forte nelle classi sociali con un reddito basso dove la Certificazione della conoscenza della lingua sembra essere associata a un miglioramento del proprio status. In effetti, le condizioni che si sono create in Grecia a causa delle politiche linguistiche adottate in questi ultimi decenni, ma anche a causa del contesto economico in cui è costretta a vivere la famiglia greca, hanno accentuato le disuguaglianze sociali. Alcune ricerche svolte in questo campo dimostrano infatti che il capitale sociale, culturale ed economico dei genitori è correlato con l'ottenimento della Certificazione di conoscenza della lingua (Tsiplakidis 2014). Questa forte attenzione alla Certificazione crea un notevole effetto *washback* durante la lezione di lingua straniera. L'ansia degli esami e la sua incidenza nella fase di apprendimento durante la lezione viene comprovata anche da studi (Cheng 2000; Shohami 2007) condotti in contesti sociali dove la Certificazione non riveste un ruolo così importante. Si può quindi facilmente intuire quanto forte possa essere questa influenza in Grecia, data la grande importanza che le viene attribuita.

Questa realtà genera pertanto una serie di problematiche. Se la Certificazione è di importanza fondamentale in Grecia per il prestigio e/o la sistemazione professionale, i test utilizzati a questo scopo dovranno assolutamente garantire:

- a. la loro adeguatezza al livello di conoscenza a cui si riferiscono,
- b. la coerenza dei contenuti nelle varie sessioni di esame,
- c. il controllo del livello di difficoltà dei loro contenuti.

La mancanza di una di queste condizioni per lo svolgimento ottimale dell'esame di conoscenza della lingua potrebbe condurre a una limitazione della validità del test e di conseguenza a un'alterazione del suo risultato: la deviazione del livello di conoscenza della lingua può condurre a un errore di misurazione e la difficoltà del contenuto agire come variabile regolatrice (e in alcuni casi come variabile parassita) dell'esame (Bachman 1990; Kane 2008). In tal caso le conseguenze sono di importanza fondamentale, dal momento che viene influenzato sia il processo decisionale a breve termine, quello rela-

tivo all'esito positivo dell'esame per il candidato, sia quello a medio e/o lungo termine, che riguarda l'uso della Certificazione da parte dello stesso candidato.

Un test di competenza linguistica comprende principalmente delle prove costituite da un foglio illustrativo che descrive le attività che il candidato deve svolgere, da uno o più testi che servono come fonte di informazione o spunto per il loro svolgimento, e da domande (items) desunte dal testo.

Il presente studio si prefigge, da una parte, di esaminare il grado di coerenza dei testi dello stesso livello linguistico che sono stati usati in diverse sessioni di esame; dall'altra, di determinare i limiti di difficoltà dei testi che possono essere usati nei livelli B1, B2, C1 e C2.

2 Metodi di misurazione della leggibilità

La necessità di una parametrizzazione delle caratteristiche che determinano la difficoltà di un testo è stata formulata fin dal 900 d.C. dai Talmudisti, i quali usarono come indice di difficoltà dei testi che studiavano il criterio delle occorrenze delle parole, oltre ad altri criteri specifici (Tekfi 1987; Zakaluk, Samuels 1996). I primi studi sistematici sulla frequenza di occorrenze nei testi scritti risalgono al 1889 e sono attribuiti al russo Nikolai Rubakin che, dopo aver analizzato 10.000 manoscritti, compilò una lista con le 1.500 parole che la maggior parte degli utenti della lingua russa doveva essere in grado di capire (Tefki 1987; Broda et al. 2014; DuBay 2007). Ricerche più sistematiche sulla leggibilità sono state condotte inizialmente da Sherman, il quale si è occupato dei parametri che rendono un testo letterario più facilmente comprensibile per il lettore, arrivando alla conclusione che la lunghezza delle frasi del testo ha una notevole influenza sulla sua leggibilità (Sherman 1893). Due decenni più tardi anche Kitson si è occupato dello stesso argomento, attraverso l'analisi di testi della stampa quotidiana e di periodici, proponendo come indici di leggibilità la lunghezza delle frasi e la lunghezza delle sillabe delle parole contenute in un testo.

Un elemento importante della ricerca di Kitson è stata la dimensione sociopsicologica da lui introdotta nello studio della leggibilità, dichiarando che importanti fattori che determinano la difficoltà di un testo sono le preferenze personali e i propri interessi (Kitson 1977). Proposte complete di formule di leggibilità appaiono dal 1923 in poi, prima tra tutte quella di Lively e Pressey, la quale ha utilizzato dati della ricerca precedentemente pubblicata da Thorndike (Bailin, Grafstein 2016). Da allora si sono succedute numerose proposte di classificazione delle formule che è possibile dividere in tre categorie principali:

- A. parametri quantitativi
- B. teorie cognitive
- C. metodi e pratiche della linguistica computazionale (Benjamin 2012; Hill, McNamara 2015; Dell’Orletta et al. 2011; Vajjala, Meurers 2014; Venturis 2018)

Le formule che appartengono alla prima categoria si basano su misurazioni quantitative del contenuto del testo da analizzare. Tali misurazioni possono essere il numero di sillabe per parola, il numero di parole per frase, e il numero di frasi per periodo. Questo tipo di indici di leggibilità ha conosciuto un vasto consenso durante l’ultimo secolo ma ha ricevuto anche delle critiche, in particolare per aver trascurato le dimensioni semantiche e sintattiche del testo, elementi fondamentali per la comprensione da parte del lettore (Davison, Kantor 1982; Vajjala, Meurers 2014). Pertanto, una formula quantitativa procede nel calcolo della leggibilità di un testo persino se contiene parole inesistenti o i suoi elementi non hanno coerenza e coesione. Le formule più conosciute di questo tipo sono Dale-Chall, Flesh e Fry. Negli ultimi anni si è tentato di affrontare questi punti deboli, introducendo ulteriori variabili e banche dati di parole. In particolare, sono state usate variabili come l’occorrenza delle parole contenute nel testo e una serie di criteri cognitivi (Chall, Dale 1995). Le formule più impiegate con queste caratteristiche sono New Dale/Chall Redability Formula, LEXILE, ATOS, Flesch-Kincaid e Read-X. I testi in lingua italiana possono essere elaborati dalle formule Franchina-Vacca e GULPEASE (Ciobanu, Dinu 2014; Garais 2011).

Sulla base delle teorie cognitive sono state sviluppate formule che misurano la leggibilità, usando variabili legate alla coerenza del testo ma anche analizzando la relazione delle sue varie componenti (McNamara, Kintsch 1996). Questa proposta ha conosciuto un grande consenso tra i ricercatori del settore. Nella sua applicazione però ha presentato difficoltà pratiche che riguardano soprattutto l’associazione dei processi cognitivi alla comprensione di un testo. Per questo motivo sono stati sviluppati processi tecnologici automatici come LSA (Latent Semantic Analysis) che rappresenta il contenuto del testo in formato vettoriale nel campo semantico (Benjamin 2012). In questa categoria vengono classificate le formule Coh-Metrix, WordNet e DeLite.

Con l’utilizzo di formule di linguistica computazionale è possibile classificare un testo a un certo livello di leggibilità sulla base di un modello di apprendimento automatico, creato con uno strumento di elaborazione del linguaggio naturale (Natural Language Processing). Tale strumento consente l’individuazione o la composizione di un insieme di criteri di classificazione, i quali possono costituire modelli linguistici, caratteristiche sintattiche, elementi di coerenza e coesione, elementi morfologici e caratteristiche semantiche dei linguaggi settoriali, aspetti del genere testuale, caratteristiche di pro-

venienza cognitiva e dati risultanti dalla ricerca sull'apprendimento della lingua seconda (Collins-Thompson, Callan 2004; Feng, Elhadad, Huenerfauth 2009; Vajjala, Meurers 2014). Nell'ambito di questa pratica furono sviluppati due tipi di analisi, SLM (Statistical Language Modeling) e SVM (Support Vector Machine). Il primo mira all'individuazione delle caratteristiche che stabiliscono la normalità del linguaggio naturale, in modo da ricercarle nei testi analizzati e attribuire il livello di leggibilità con maggiore fedeltà (Rosenfield 2000; Si, Callan 2001). Il secondo mira alla classificazione di una serie di testi sulla base di dati addestrati (*trained data*) che sono stati creati in una fase precedente (Joachims 2000; Vapnik 1982).

3 Caratteristiche metodologiche della ricerca

L'ipotesi di ricerca primaria di questo studio è che il livello linguistico dei testi possa essere identificato in base all'indice della loro leggibilità. Partendo da questa ipotesi si è cercato di indagare i limiti di leggibilità dei testi che permettono la loro categorizzazione in uno dei livelli B1, B2, C1, C2. Un'ipotesi di ricerca secondaria è che i criteri quantitativi-superficiali non costituiscano un chiaro indice del livello linguistico del testo. Per la verifica di questa ipotesi è stato controllato il grado di pertinenza della lunghezza testuale, frastica e lessicale con il livello di conoscenza della lingua. Un'ulteriore ipotesi di ricerca è che nel definire il grado di difficoltà dei testi il contributo delle variabili sintattiche e lessicali sia maggiore delle variabili superficiali (quantitative).

La ricerca si è basata su un corpus di 184 testi raccolti dall'insieme dei testi usati negli esami di Certificazione di Conoscenza della Lingua Italiana della Certificazione Statale Greca per testare la comprensione scritta (unità 1), nel corso degli anni 2011-16 (12 sessioni d'esame). La scelta di questo periodo si deve al fatto che dal 2011 è stato introdotto il test graduato nel sistema della Certificazione Statale di conoscenza della lingua (KPG), per cui si prevede di arrivare a una chiara distinzione tra i testi di diverso livello che coesistono nello stesso test.

Il KPG è un sistema di certificazione di conoscenza delle lingue straniere, glocal e multilingue, sviluppato in base ai principi del Quadro Comune Europeo e diretto dal Ministero della Pubblica Istruzione della Grecia. Opera dal 2003 e organizza esami per la certificazione di 6 lingue (francese, inglese, italiano, spagnolo, tedesco, turco) per tutti i livelli del QCER (A, B, C). Una delle particolarità del KPG è che le prove sono strutturate in 3 livelli che all'interno presentano prove graduate (A: A1-A2, B: B1-B2, C: C1-C2). Un'altra particolarità importante del sistema è che per i livelli B e C comprende anche una prova di mediazione, come prevede il QCER, sia nella fase della pro-

duzione orale che della produzione scritta. La responsabilità scientifica del sistema è stata affidata alle Università Aristotele di Salonicco e Capodistriaca di Atene e in particolare a un gruppo di professori esperti in testing e valutazione linguistica (Dendrinou, Karavas 2013).

La distribuzione dei testi nelle sessioni di esame inclusi nel campione si può osservare nella tabella 1:

Tabella 1 Distribuzione dei testi per sessione d'esame e livello

		Livello				Totale
		B1	B2	C1	C2	
Sessione	2011A	4	4	4	0	12
	2011B	5	4	6	0	15
	2012A	4	5	5	0	14
	2012B	4	4	7	0	15
	2013A	4	4	7	0	15
	2013B	4	4	5	3	16
	2014A	4	5	4	4	17
	2014B	4	4	5	3	16
	2015A	5	4	5	3	17
	2015B	4	4	4	4	16
	2016A	3	5	4	4	16
2016B	3	4	4	4	15	
Totale		48	51	60	25	184

Il livello linguistico dei testi compresi nel corpus è stato fissato dagli autori che fanno parte della commissione per la preparazione delle prove per gli esami di lingua italiana della Certificazione statale, in base alle specifiche del sistema (National and Kapodistrian University of Athens, Aristotle University of Thessaloniki 2016).

Per l'analisi dei testi è stato usato il software DyLan TextTools v2.1.9,¹ il quale implementa l'analisi e la misurazione della leggibilità dei testi di lingua italiana e inglese, combinando criteri quantitativi con criteri cognitivi e sfruttando tecniche di linguistica computazionale. Il risultato finale è la classificazione di ogni testo analizzato in un certo grado di difficoltà, determinato numericamente nella scala da 0 a 100. Questo software è in realtà SVM il quale usa una libreria di testi (LISVM). Sulla base di una serie di caratteristiche del testo analizzato, ma anche di un corpus addestrato di testi, crea un modello statistico e misura con questo la leggibilità del testo introdotto nel campo di analisi (Dell'Orletta et al. 2011). Tra l'altro questo pro-

¹ Nella letteratura pertinente questo software viene indicato come READ-IT, ma nella sua versione online ha il nome menzionato all'inizio del paragrafo.

gramma misura la leggibilità con la formula GULPEASE, consentendo di confrontare il risultato ottenuto con quello di una formula basata su caratteristiche quantitative del testo analizzato.

Nell'analisi condotta sono state comprese 17 variabili, 4 delle quali erano categoriche-qualitative e le restanti quantitative [tab. 2]. Le variabili categoriche comprendevano la sessione d'esame, il livello, la prova degli esami in cui è stato usato il testo (*task*), la coesione del testo come elemento degli esami, vale a dire se il testo era continuo, conteneva spazi o era scomposto in paragrafi senza coerenza. Le variabili quantitative erano gli indici di READ-IT (variabili 5-8), delle quali READ-IT Global riguarda il valore complessivo di leggibilità. READ-IT Base è il valore di leggibilità sulla base delle caratteristiche quantitative del testo, ossia, della media di parole per frase e della media dei caratteri per parola. L'indice READ-IT Lexical si basa sulle caratteristiche lessicali e morfo-sintattiche del testo. I parametri presi in considerazione per il calcolo di questo indice sono la distribuzione del vocabolario testuale secondo le categorie di frequenza d'uso, definite da De Mauro (2000), il rapporto tra i tipi di parole singole (*types*) e il numero delle parole del testo (*tokens*), il modello di probabilità linguistica delle parti monogrammate del discorso (il modello linguistico che calcola la probabilità che una parola appaia in un testo), la densità lessicale (rapporto tra parole complete e la totalità delle parole) e la distribuzione dei verbi (distribuzione dei verbi secondo la loro disponibilità) (Dell'Orletta et al. 2011).

L'indice READ-IT Syntactic si forma sulla base dei parametri sintattici della probabilità anormale di dipendenza dei tipi sintattici (es. soggetto, oggetto diretto, definizione), della profondità dei componenti dell'albero di analisi, degli elementi della categoria verbale, degli elementi di subordinazione e della gamma dei dati sulla coesione (Feng et al. 2010; Dell'Orletta et al. 2011; 2014).

Tabella 2 Variabili della ricerca

Statistics		
Valid	N	Missing
1. Sessione	184	0
2. Livello	184	0
3. Task	184	0
4. Coesione	184	0
5. READ-IT Base	184	0
6. READ-IT Lexical	184	0
7. READ-IT Syntactic	184	0
8. READ-IT Global	184	0
9. GULPEASE	184	0
10. Tipo/Unità	184	0
11. Densità lessicale	184	0
12. Vocabolario di base	184	0
13. VdB Fondamentale	184	0
14. VdB Alto Uso	184	0
15. VdB Alta disponibilità	184	0
16. Proposizioni principali	184	0
17. Proposizioni subordinanti	184	0

La nona variabile riguarda l'indice di leggibilità GULPEASE (Tonelli, Manh, Pianta 2012), mentre le variabili 10-17 costituiscono singoli elementi che concorrono alla formazione del risultato dell'indice sintattico o lessicale, che si è ritenuto opportuno analizzare ulteriormente in modo da valutare il loro eventuale particolare contributo alla formazione del risultato finale di leggibilità.

L'analisi dei dati raccolti utilizzando gli strumenti di testo DyLan v2.1.9 è stata eseguita con l'aiuto del software SPSS 24.

4 Risultati della ricerca

Dall'elaborazione iniziale dei dati raccolti è emerso che mentre l'indice di ranking READ-IT Global evidenzia la differenziazione dei testi in base al livello di conoscenza della lingua per il quale è stato selezionato, l'indice GULPEASE, al contrario, non lo raggiunge.

Tabella 3 Media di difficoltà degli indici della ricerca e deviazione

		Statistica Descrittiva		
Level		Mean	Std. Deviation	N
B1	READ-IT Base	46.228	26.7397	48
	READ-IT Lexical	68.483	33.8988	48
	READ-IT Syntactic	64.040	32.9941	48
	READ-IT Global	59.585	39.2900	48
	GULPEASE	57.200	<u>8.4120</u>	48
B2	READ-IT Base	41.471	30.5911	51
	READ-IT Lexical	71.104	32.3373	51
	READ-IT Syntactic	67.493	33.6823	51
	READ-IT Global	72.284	34.3752	51
	GULPEASE	59.616	<u>11.8003</u>	51
C1	READ-IT Base	55.197	33.1812	60
	READ-IT Lexical	78.910	29.2616	60
	READ-IT Syntactic	76.393	26.7880	60
	READ-IT Global	87.910	19.2659	60
	GULPEASE	55.170	<u>11.2533</u>	60
C2	READ-IT Base	66.048	31.8251	25
	READ-IT Lexical	87.432	22.0423	25
	READ-IT Syntactic	76.128	33.6953	25
	READ-IT Global	94.512	20.1307	25
	GULPEASE	54.972	<u>15.1254</u>	25

In particolare, nella tabella 3 troviamo che il grado medio di difficoltà dei testi analizzati è 59.585 per il livello B1, 72.284 per il livello B2, 87.910 per il livello C1 e 94.512 per il livello C2.

La graduazione della difficoltà sembra essere in linea con la differenza di difficoltà risultante dalle descrizioni del Quadro Comune Europeo di Riferimento (Consiglio d'Europa 2001). Osservando però la colonna della deviazione standard, si nota che esiste una differenza significativa nel grado di leggibilità dei testi utilizzati negli esami del KPG: l'indice READ-IT Global presenta SD = 39,29 mentre tutti e tre gli indicatori dai quali risulta la leggibilità totale mostrano una deviazione standard notevolmente elevata.

Questo fatto conferma l'ipotesi della ricerca secondo la quale i criteri quantitativi per la selezione di un testo, come la quantità di parole, combinati ad altri elementi come ad esempio l'argomento, non sono sufficienti per una scelta adeguata al livello prestabilito di co-

noscenza della lingua; soprattutto per una scelta coerente con le caratteristiche dei testi inclusi per lo stesso scopo negli anni precedenti e valutati in base al loro livello di difficoltà. A questa conclusione si arriva anche per il fatto che la deviazione standard dell'indice del criterio di superficie GULPEASE è molto più piccola, dando l'impressione che i testi selezionati per ogni livello in periodi diversi di esame non differiscano molto tra loro nei livelli di difficoltà. Questa conclusione però viene ribaltata applicando una formula che utilizza criteri diversi come, in questo caso, READ-IT.

Il controllo di correlazione tra i risultati degli indicatori finali (READ-IT Global, GULPEASE) e intermedi (READ-IT Base, READ-IT Lexical, READ-IT Syntactic) ha guidato alle seguenti informazioni:

Per il livello B1 esiste una forte correlazione (livello di significatività 0,01) dell'indice lessicale con il READ-IT finale (0,538). Lo stesso fenomeno si verifica al livello B2 (0,593). Al livello C l'indice che risulta essere fortemente correlato al risultato finale dell'indice finale di leggibilità (READ-IT finale) è quello delle caratteristiche sintattiche del testo (rispettivamente READ-IT Syntactic), mentre solo per il livello C1 appare come coerente (0,430) al risultato finale della difficoltà del testo anche l'indice delle caratteristiche quantitative (READ-IT Base). Da questi dati possiamo concludere che a livello B (B1 e B2) le variabili lessicali e morfosintattiche che compongono l'indice READ-IT Lexical (es. densità lessicale) portano un contributo particolarmente importante nella rappresentazione della difficoltà dei testi selezionati per questo livello, mentre per il livello C il contributo delle variabili sintattiche sembra essere più rilevante, tale da determinare in sostanza il risultato finale di leggibilità. La relazione delle variabili di estensione con il grado di difficoltà del livello C1 può sembrare interpretabile per via della relazione tra la profondità delle componenti dell'albero di analisi, che conta, tra gli altri, l'indice sintattico della ricerca e il numero di parole della frase misurate dall'indice READ-IT. Nei testi di livello C2, tuttavia, la coerenza delle variabili sintattiche con il risultato finale di leggibilità risulta più evidente, dato che variabili come la probabilità anormale di dipendenza dei tipi sintattici, rendono più evidente la loro presenza.

Per quanto riguarda l'indice GULPEASE, si osserva che al livello C ha un'ampia correlazione con l'indice quantitativo READ-IT ma in rapporto inverso (all'aumentare di un valore diminuisce l'altro). Questo fatto potrebbe sollevare dubbi se non fosse per il fatto che l'indice GULPEASE non riesce a distinguere il livello di conoscenza della lingua dei testi in base alla loro difficoltà. Mostra infatti i testi di livello B più difficili di quelli di livello C, fatto che difficilmente potrebbe essere accettato, se si tiene conto delle loro caratteristiche lessicali e sintattiche.

Questa interpretazione viene probabilmente confermata in quanto in entrambi i livelli compare una relazione inversamente propor-

zionale dello stesso indice con la sintassi (C1 e C2), mentre al livello C1 è correlato anche con l'indice READ-IT complessivo.

Tabella 4 Correlazione tra indici di leggibilità finali e intermedi (livello B)

Level			READ-IT Base	READ-IT Lexical	READ-IT Syntactic	READ-IT Global	GULPEASE	
B1	READ-IT Base	Pearson Cor	1	.017	.259	-.110	-.021	
		Sig. (2-tailed)		.909	.076	.459	.890	
		N	48	48	48	48	48	
	READ-IT Lexical	Pearson Cor	.017	1	-.085	.538**	-.203	
		Sig. (2-tailed)	.909		.564	.000	.171	
		N	48	48	48	48	48	
	READ-IT Syntactic	Pearson Cor	.259	-.085	1	.122	-.006	
		Sig. (2-tailed)	.076	.564		.407	.970	
		N	48	48	48	48	48	
	READ-IT Global	Pearson Cor	-.110	.538**	.122	1	1	
		Sig. (2-tailed)	.459	.000	.407			
		N	48	48	48	48	48	
	GULPEASE	Pearson Cor	-.021	-.203	-.006	1	.361*	
		Sig. (2-tailed)	.890	.171	.970		.013	
		N	48	48	48	48	48	
	B2	READ-IT Base	Pearson Cor	1	.050	.581**	.163	.186
			Sig. (2-tailed)		.728	.000	.254	.222
			N	51	51	51	51	51
READ-IT Lexical		Pearson Cor	.050	1	.118	.593**	-.432	
		Sig. (2-tailed)	.728		.409	.000	.003	
		N	51	51	51	51	51	
READ-IT Syntactic		Pearson Cor	.581**	.118	1	.222	-.219	
		Sig. (2-tailed)	.000	.409		.118	.148	
		N	51	51	51	51	51	
READ-IT Global		Pearson Cor	.163	.593**	.222	1	1	
		Sig. (2-tailed)	.254	.000	.118			
		N	51	51	51	51	51	
GULPEASE		Pearson Cor	.186	-.432	-.219	1	.488	
		Sig. (2-tailed)	.222	.003	.148		.001	
		N	51	51	51	51	51	

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Tabella 5 Correlazione tra indici di leggibilità finali e intermedi (livello C)

Level			READ-IT Base	READ-IT Lexical	READ-IT Syntactic	READ-IT Global	GULPEASE	
C1	READ-IT Base	Pearson Cor	1	-.210	.560**	.430**	-.804**	
		Sig. (2-tailed)		.107	.000	.001	.000	
		N	60	60	60	60	60	
	READ-IT Lexical	Pearson Cor	-.210	1	-.168	.131	.216	
		Sig. (2-tailed)	.107		.198	.318	.097	
		N	60	60	60	60	60	
	READ-IT Syntactic	Pearson Cor	.560**	-.168	1	.476**	-.504**	
		Sig. (2-tailed)	.000	.198		.000	.000	
		N	60	60	60	60	60	
	READ-IT Global	Pearson Cor	.430**	.131	.476**	1	-.332**	
		Sig. (2-tailed)	.001	.318	.000		.010	
		N	60	60	60	60	60	
	GULPEASE	Pearson Cor	-.804**	.216	-.504**	-.332**	1	
		Sig. (2-tailed)	.000	.097	.000	.010		
		N	60	60	60	60	60	
	C2	READ-IT Base	Pearson Cor	1	-.273	.536**	.052	-.904**
			Sig. (2-tailed)		.187	.006	.804	.000
			N	25	25	25	25	25
READ-IT Lexical		Pearson Cor	-.273	1	-.347	-.146	.262	
		Sig. (2-tailed)	.187		.090	.485	.205	
		N	25	25	25	25	25	
READ-IT Syntactic		Pearson Cor	.536**	-.347	1	.506**	-.570**	
		Sig. (2-tailed)	.006	.090		.010	.003	
		N	25	25	25	25	25	
READ-IT Global		Pearson Cor	.052	-.146	.560**	1	-.025	
		Sig. (2-tailed)	.804	.485	.010		.906	
		N	25	25	25	25	25	
GULPEASE		Pearson Cor	-.904**	.262	-.570**	-.025	1	
		Sig. (2-tailed)	.000	.205	.003	.906		
		N	25	25	25	25	25	

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

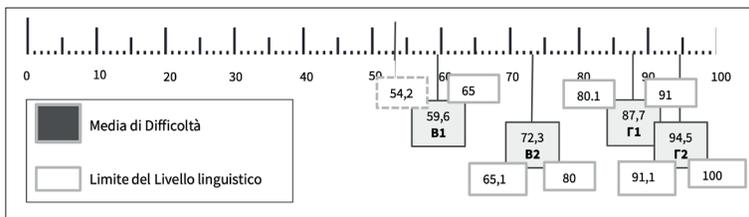
5 Conclusioni

Dal trattamento dei dati raccolti e presentati risulta che la selezione di un testo per uso valutativo o didattico, basata esclusivamente sulle caratteristiche del suo contenuto, non sembra consentire la scelta appropriata.

Alla stessa conclusione si arriva anche per quanto riguarda la coerenza della selezione per il livello di difficoltà del testo, in quanto tali informazioni possono fornire dati basati su variabili lessicali, morfosintattiche e sintattiche.

Pertanto, una formula che utilizza tali criteri per calcolare il grado di leggibilità di un testo può essere un utile supporto per un autore di test linguistici di comprensione di un testo o di attività create per uso didattico.

Un'altra considerazione da fare è che può esistere una corrispondenza tra il grado di difficoltà di un testo e il suo livello di competenza linguistica. Sulla base dei dati presentati in precedenza vengono proposti i seguenti limiti di difficoltà di un testo di lingua italiana per ogni livello di conoscenza della lingua.



I limiti di difficoltà dei livelli derivano dalla formula:

$$\lim_{i \rightarrow j} RG = (X_{jmax} - X_{imin}) : 2$$

dove $\lim_{i \rightarrow j} RG$ simbolizza il limite di transizione dal valore READ-IT Global di ogni livello di conoscenza della lingua successivo (i, j) X_{jmax} a quello immediatamente successivo (i→j), il valore più alto di ogni livello di conoscenza della lingua successivo (i, j) e X_{imin} il valore più basso di ogni livello di conoscenza della lingua successivo.

Va sottolineato che la scala formata sulla base di quanto sopra è a intervalli; indica cioè una relazione gerarchica tra i suoi gradi, con intervalli conosciuti ma non uguali.

Poiché mancano i dati sui livelli A1 e A2 che non interessano il presente studio, la scala risulta incompleta e di conseguenza risulta impossibile calcolare il livello inferiore al B1.

Per il momento e fino al completamento della scala, si potrebbe utilizzare come limite inferiore del livello B1, la gamma della parte superiore della classe, ovvero 5.4. In questo caso il limite inferiore viene fissato a 53,3.

In conclusione, determinare il livello di difficoltà di un testo utilizzato nel quadro di un processo di valutazione di conoscenza della lingua, ma anche dell'insegnamento delle lingue, è un'esigenza che, se non soddisfatta, rischia di creare condizioni che potrebbero portare a scelte inappropriate e inaffidabili nell'ambiente sociale.

La pratica migliore sarebbe ovviamente quella di compilare limiti standard, sulla base di testi utilizzati da tutti i sistemi validi di certificazione della lingua italiana e tenendo in considerazione il corpus dei testi provenienti da ricerche pertinenti (Baroni et al. 2009) in cui sono state create delle categorie in base al livello di conoscenza della lingua. Una tale selezione sarà soggetta a limitazioni in quanto la difficoltà di un testo è legata anche alle caratteristiche del lettore (esperienze, background socioculturale), che differiscono a seconda del suo contesto sociale, economico e/o operativo.

Bibliografia

- Bachman, L.F. (1990). *Fundamental Considerations in Language Testing*. Oxford; New York: Oxford University Press. <https://doi.org/10.1017/CBO9781107415324.004>.
- Bailin, A.; Grafstein, A. (2016). «Readability: Text and Context». *Readability: Text and Context*, January, 1-224. <https://doi.org/10.1057/9781137388773>.
- Baroni, M.; Bernardini, S.; Ferraresi, A.; Zanchetta, E. (2009). «The WaCky wide web: A Collection of Very Large Linguistically Processed Web-Crawled Corpora». *Language Resources and Evaluation*, 43(3), 209-26. <https://doi.org/10.1007/s10579-009-9081-4>.
- Benjamin, R.G. (2012). «Reconstructing Readability: Recent Developments and Recommendations in the Analysis of Text Difficulty». *Educational Psychology Review*, 24(1), 63-88. <https://doi.org/10.1007/s10648-011-9181-8>.
- Broda, B.; Ogrodniczuk, M.; Niton, B.; Gruszczyński, W. (2014). «Measuring Readability of Polish Texts: Baseline Experiments». *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 573-80. <http://nlp.ipipan.waw.pl/Bib/broda:etal:14:lrec.pdf>.
- Chall, J.; Dale, E. (1995). *Readability Revisited: The New Dale-Chall Readability Formula*. Cambridge: Brookline Books.
- Cheng, L.; Curtis, A. (2004). «Washback or Backwash: A Review of the Impact of Testing on Teaching and Learning». *Washback in Language Testing: Research Contexts and Methods*, 27, 3-17.
- Ciobanu, A.M.; Dinu, L. (2014). «A Quantitative Insight into the Impact of Translation on Readability». Williams, S.; Siddharthan, A.; Nenkova, A. (eds), *Proceedings of the 3rd Workshop on Predicting and Improving Text Readability for Target Reader Populations (PITR)*. Gothenburg: Association for Computational Linguistics, 104-13. <http://dx.doi.org/10.3115/v1/W14-1212>.

- Collins-Thompson, K.; Callan, J. (2004). «A Language Modeling Approach to Predicting Reading Difficulty». *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL 2004)*. Boston (MA): Association for Computational Linguistics, 193-200. http://acl.ldc.upenn.edu/hlt-naacl2004/main/pdf/111_Paper.pdf.
- Council of Europe. (2001). *The Common European Framework in Its Political and Educational Context: Learning, teaching, assessment*. Strasbourg: Cambridge University Press. <https://rm.coe.int/1680459f97>.
- Davison, A.; Kantor, R. (1982). «On the Failure of Readability Formulas to Define Readable Texts: A Case Study from Adaptations». *Reading Research Quarterly*, 17(2), 187-209.
- Dell'Orletta, F.; Montemagni, S.; Venturi, G.; Zampolli, A.; Moruzzi, G. (2011). «READ-IT: Assessing Readability of Italian Texts with a View to Text Simplification». *Computational Linguistics*, 73-83.
- Dell'Orletta, F.; Wieling, M.; Cimino, A.; Venturi, G.; Montemagni, S. (2014). «Assessing the Readability of Sentences: Which Corpora and Features?». Alm, N. (ed.), *Proceedings of the Ninth Workshop on Innovative Use of NLP for Building Educational Applications*. Edinburgh: Association for Computational Linguistics, 163-73. <http://www.aclweb.org/anthology/W14/W14-1820>.
- De Mauro, T. (2000). *Il dizionario della lingua italiana*. Torino: Paravia.
- Dendrinis, B.; Karavas, K. (2013). *The Greek Foreign Language Examinations for the State Certificate of Language Proficiency. The KPG Handbook. Performance Descriptors and Specifications*. Athens: RCEL publications. https://rceL2.enl.uoa.gr/kpg/files/KPG_Handbook_17X24.pdf.
- Dendrinis, B.; Zouganeli, K.; Karavas, K. (2013). *Foreign Language Learning. European Survey on Language Competences*. Athens: National and Kapodistrian University of Athens. http://www.rcel.enl.uoa.gr/fileadmin/elp.enl.uoa.gr/uploads/ESLC_EN_WEB.pdf.
- DuBay, W.H. (2007). *Smart Language. Readers, Readability, and the Grading of Text*. Costa Mesa (CA): Impact Information.
- Feng, L.; Elhadad, N.; Huenerfauth, M. (2009). «Cognitively Motivated Features for Readability Assessment». *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*. Athens: The Association for Computational Linguistics, 229-37.
- Feng, L.; Jansche, M.; Huenerfauth, M.; Elhadad, N. (2010). «A Comparison of Features for Automatic Readability Assessment». *COLIN '10: Proceedings of the 23rd International Conference on Computational Linguistics: Posters* (August). Stroudsburg (PA): Association for Computational Linguistics, 276-84. <https://dl.acm.org/citation.cfm?id=1944598>.
- Garais, E.-G. (2011). «Web Applications Readability». *Journal of Information Systems and Operations Management*, 5(1), 117-21.
- Hill, K.; McNamara, T. (2015). «Validity Inferences Under High-Stakes Conditions: A Response From Language Testing». *Measurement*, 13(1), 39-43. <https://doi.org/10.1080/15366367.2015.1016334>.
- Joachims, T. (2000). *Estimating the Generalization Performance of a SVM Efficiently*. Dortmund: Universität Dortmund.
- Kane, M. (2008). «Errors of Measurement, Theory, and Public Policy». *Policy*, 40. <http://www.ets.org/Media/Research/pdf/PICANG12.pdf>.

- Kitson, H.D. (1927). *The Mind Of The Buyer-A-Psychology Of Selling*. New York: The Macmillan Company.
- Mattheoudakis, M.; Alexiou, T. (2009). «Early Foreign Language Instruction in Greece: Socioeconomic Factors and their Effect on Young Learners' Language Development». *The Age Factor and Early Language Learning*, 40, 227-52. <https://doi.org/10.1515/9783110218282.227>.
- McNamara, D.S.; Kintsch, W. (1996). «Learning from Texts: Effects of Prior Knowledge and Text Coherence». *Discourse Processes*, 22(3), 247-88.
- National and Kapodistrian University of Athens; Aristotle University of Thessaloniki (2016). Προδιαγραφές του συστήματος αξιολόγησης για το ΚΠΓ (Standards del sistema di valutazione per il KPG). Athens: RCEl. http://rce12.enl.uoa.gr/kpg/files/KPG_Prodiagrafes.pdf.
- Psitos, M.; Marotta, C. (2018). «Tratti neo-standard nel parlato degli insegnanti di italiano LS: uno studio a Salonicco». *Italiano LinguaDue*, 9(2), 97-120. <https://riviste.unimi.it/index.php/promoitals/article/view/9873/9345>.
- Rosenfield, R. (2000). «Two Decades of Statistical Language Modeling: Where Do We Go from Here?». *Proceedings of the IEEE*, 88(8), 1270-8.
- Sherman, L.A. (1893). *Analytics of Literature: A Manual for the Objective Study of English Prose and Poetry*. Ginn & Co. Boston: Ginn & Company. <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:N+Title#0>.
- Shohamy, E. (2007). «Language Tests as Language Policy Tools». *Assessment in Education: Principles, Policy & Practice*, 14(1), 117-30. <https://doi.org/10.1080/09695940701272948>.
- Si, L.; Callan, J. (2001). «A Statistical Model for Scientific Readability». *Proceedings IEEE 2000 First International Symposium on Quality Electronic Design* (San Jose, California, 20-22 March 2000). Los Alamitos (CA): IEE Computer Society, 303-7. <https://doi.org/10.1109/ISQED.2000.838852>.
- Tekfi, C. (1987). «Readability Formulas: An Overview». *Journal of Documentation*, 43(3), 261-73.
- Tonelli, S.; Manh, K.T.; Pianta, E. (2012). «Making Readability Indices Readable». *Proceedings of the First Workshop on Predicting and Improving Text Readability for Target Reader Populations* (Montréal, Canada, 7 June 2012). Stroudsburg (PA): Association for Computational Linguistics, 40-8. <https://dl.acm.org/citation.cfm?id=2390924>.
- Tsiplakidis, I. (2014). Το μάθημα της αγγλικής γλώσσας στη δημόσια Δευτεροβάθμια Εκπαίδευση και η προσφυγή στο φροντιστήριο: μια εμπειρική έρευνα για την επίδραση της κοινωνικής τάξης στην εκμάθηση και πιστοποίησή της (La materia della lingua inglese nella Pubblica Istruzione secondaria e il ricorso agli istituti: Una ricerca empirica sull'influenza della classe sociale sul suo apprendimento e certificazione) [PhD Dissertation]. Ioannina: University of Ioannina. <https://www.didaktorika.gr/eadd/handle/10442/39561>.
- Vajjala, S.; Meurers, D. (2014). «Readability Assessment for Text Simplification: From Analyzing Documents to Identifying Sentential Simplifications». *International Journal of Applied Linguistics*, 165(2), 1-23. <https://doi.org/10.1075/itl.165.2.04vaj>.
- Vapnik, V. (1982). *Estimation of Dependences Based on Empirical Data*. 2nd ed. New York: Springer-Verlag.

- Venturis, A. (2018). «La selezione di testi italiani per il controllo della comprensione scritta: indicatori di livello linguistico e di difficoltà». Pirvu, E. (a cura di), *Il tempo e lo spazio nella lingua e nella letteratura italiana*. Firenze: Franco Cesati Editore, 243-58.
- Zakaluk, B.L.; Samuels, S.J. (1996). «Issues Related to Text Comprehensibility: The Future of Readability». *Revue Québécoise de Linguistique*, 25(1), 41-59. <https://doi.org/10.7202/603126ar>.

