Calculation of Degree of Difficulty and Measurement of Readability of Italian Texts of the KPG

Apostolos Klonis

Teacher of Informatics, Ph.D., in Software Technology, Serres, Greece

ABSTRACT

In the context of this article, the construction of software for language proficiency level and difficulty measurement of language proficiency tests was sought. The main software goal can not help with identification issues, and also the recognition of the text can not be mitigated by using common scoring systems, including the types of recognition calculations. This software accepts various Italian language test subjects whether they are graded or not, and after analyzing the text, ranks these texts at the corresponding Italian language proficiency levels and calculates the degree of difficulty of the test. The selection of subjects and their inclusion is of course done by the sampling method.

1. Introduction

The degree of ease of reading a text is interwoven with the notation of the symbols, the decipherment of the content of the words, the identification of symbols with appropriate words that semantically indicate these symbols, with the ultimate goal of penetrating the meaning of the text. Only then does the text function as a key element of learning and mental cultivation for the student. Of course, there are many factors involved in understanding the text, which are related to:
• the reader,
• the text itself,
• the environment in which the reading activity takes place.

The difficulty of reading a text may depend on factors that shape the author of the text such as:
• sentence length,
• the word length,
• paragraph length,
• and more generally the clarification of the text (clarifying text).

The first systematic studies of the frequency of words in written texts were made in 1889 by the Russian Nikolai Rubakin, who, after analyzing 10,000 manuscripts, compiled a list of 1,500 words that most users of the Russian language should be able to understand. More systematic research on readability was initially conducted by Sherman, who dealt with the parameters that make a literary text easier to understand for the reader and concluded that the length of the phrases it contains significantly affects its readability.

Two decades later, Kitson explored the same subject, based on daily and periodical texts, and suggested as readability indicators the range of phrases and the range of syllables of words contained in a text. An important element of Kitson's research was the socio-psychological dimension he introduced in the study of readability, stating that important factors in shaping the difficulty of a text are his personal preferences and interests.

Complete readability formula proposals appear from 1923 onwards, with the first being by Lively and Pressey, which used data from Thorndike's earlier research. The analysis and study of the factors that affect readability have been the subject of many linguistic and psychological researches after 1950, which led to the adoption of some direct or indirect indicators, which measure the ease or difficulty of reading and understanding from one point of view. average student. One such indicator is the degree or readability index.

By readability or Ease of Readability, (John R. Bormuth, 2006) we mean the easy, fast and correct recognition of the forms of letters, words and symbols of a text that implies its comprehension and decoding. The particular characteristics of each text have led many linguists to formulate mathematical formulas that link them and constitute an objective way of calculating / estimating the readability of texts.

Several years ago until today there are several formulas for measuring text difficulty, such as formulas that use traditional readability criteria, formulas based on cognitive theories, formulas based on statistical analysis and mixed methods. This article presents the results of our research after using the software used.

Tele:
E-mail address: apoklonis@gmail.com

2. Research Methodology

During the experimental application of the software, a field research was conducted, during which not only the active students of AUTh. but also many other external internet users to evaluate the software. Its evaluation was carried out in two phases. The first was immediately after the completion of the software and the second after its corrections and any improvements. Additional features of the users were also investigated, as the effectiveness of the program may be influenced by other variables, such as age, gender, education and more generally the educational level of the users.

In addition, there were closed-ended issues, which facilitated the processing and analysis of the data collected.

The software developed is Test Calculator. Test Calculator is a software that shows us the level of language proficiency of Italian texts and measures the difficulty of tests in an Italian text. For the analysis of the collected data we used the service SurveyMonkey (Knussen & McFadyen, 2010), which is provided free of charge via the internet and understands very well the special role of the questionnaires.

This was chosen from the many sites that offer similar services, because it is simple to create questionnaires, does not provide difficulties in developing many relevant questions and is not mentioned somewhere vaguely (Waclawski, 2012). The variables given are understandable and help the reader to understand them more and to obtain as much information as possible.

Specifically, the software created for this purpose provides the following functions:

● The reader can enter his text and his software extracts the results and scores from the above types of readability as well as the level of difficulty of the tests.

• It can also see some additional functions of the system such as word counts, syllables, characters, sentences, characters per word, syllables per word, words per sentence.

• There is also the ability to read text on an entire web page, the user typing the corresponding address, and the ability to measure the readability of .doc, .docx, and .pdf files, if these files are uploaded to the software without formatting.

Specifically at this stage, however, concerted efforts were made to consolidate the criteria for determining the level of difficulty of the written and oral comprehension tests, as well as the criteria for determining the level of difficulty of the written and oral production tests for the Italian language. Once the texts are entered in the software, after processing and analysis of these language tests, they are classified in the respective language levels and in the respective degrees of difficulty of the tests, after first calculating the degree of difficulty in the whole text.Detailed descriptions of the software and calculations of the various text readability indicators will be analyzed below.

This new software sought to be created is a more reliable and valid solution than other software available freely on the Internet to assess the difficulty of Italian language tests and is used in the development of test questions for the various levels of language proficiency certification. It is the only one with these functions and these possibilities. In addition, this software uses the Italian dictionary as a database, thus checking only those words that it supports and "understands". All other words are ignored.

Finally, the reader now has at his disposal a readability software which:

a) uses more parameters with increased discretion between language levels;

b) can give a direct correspondence between the degree of readability and the level of language proficiency;

c) calculates the probabilities corresponding to all levels of language learning and their degree of difficulty.

3. Results

In this research we took the topics of previous years from the State Certificate of Language Proficiency level B (BI & B2) for the Italian language and entered in our software (Test Calculator) the following text, to measure the level of language proficiency and the degree of difficulty specific test.
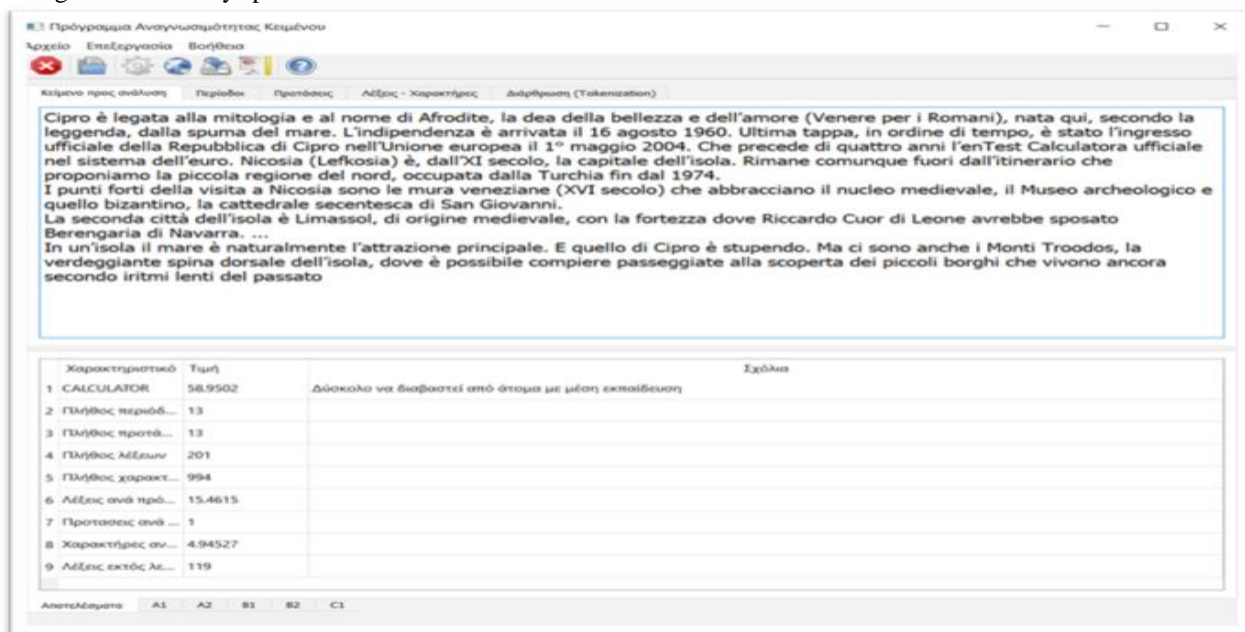
Figure 1. The Test Calculator software after entering the Italian text.

«Cipro è legata alla mitologia e al nome di Afrodite, la dea della bellezza e dell'amore (Venere per i Romani), nata qui, secondo la leggenda, dalla spuma del mare. L'indipendenza è arrivata il 16 agosto 1960. Ultima tappa, in ordine di tempo, è stato l'ingresso ufficiale della Repubblica di Cipro nell'Unione europea il 1° maggio 2004. Che precede di quattro anni

l'enTest Calculatora ufficiale nel sistema dell'euro. Nicosia (Lefkosia) è, dall'XI secolo, la capitale dell'isola. Rimane comunque fuori dall'itinerario che proponiamo la piccola regione del nord, occupata dalla Turchia fin dal 1974.

I punti forti della visita a Nicosia sono le mura veneziane (XVI secolo) che abbracciano il nucleo medievale, il Museo archeologico e quello bizantino, la cattedrale secentesca di San Giovanni.

La seconda città dell'isola è Limassol, di origine medievale, con la fortezza dove Riccardo Cuor di Leone avrebbe sposato Berengaria di Navarra. ...

In un'isola il mare è naturalmente l'attrazione principale. E quello di Cipro è stupendo. Ma ci sono anche i Monti Troodos, la verdeggiante spina dorsale dell'isola, dove è possibile compiere passeggiate alla scoperta dei piccoli borghi che vivono ancora secondo iritmi lenti del passato «

After the execution of the program, we see the following results of the features with their corresponding values in the right column and next to the comments from the execution of our program.

	Χαρακτηριστικό	Τιμή	Σχόλια
1	CALCULATOR	58.9502	Δύσκολο να διαβαστεί από άτομα με μέση εκπαίδευση
2	Πλήθος περιόδ...	13	
3	Πλήθος προτά...	13	
4	Πλήθος λέξεων	201	
5	Πλήθος χαρακτ...	994	
6	Λέξεις ανά πρό...	15.4615	
7	Προτασεις ανά ...	1	
8	Χαρακτήρες αν...	4.94527	
9	Λέξεις εκτός λε...	119	

Figure 2. The results of the features after the introduction of the Italian text.

Thus, after the calculation we have the following results of the variables with their respective values:

Calculator: 58,95
Πλήθος περιόδων: 13
Πλήθος προτάσεων: 13
Πλήθος λέξεων: 201
Πλήθος χαρακτήρων: 994
Λέξεις ανά πρόταση: 15.4615
Προτάσεις ανά περίοδο: 1
Χαρακτήρες ανά λέξη: 4,94
Λέξεις εκτός λεξικών: 119
Σχόλια: Δύσκολο να διαβαστεί από άτομα με μέση εκπαίδευση.

Κείμενο προς ανάλυση	Περίοδοι	Προτάσεις	Λέξεις - Χαρακτήρες	Διάρθρωση (Tokenization)

	Περίοδοι
1	Cipro è legata alla mitologia e al nome di Afrodite, la dea della bellezza e dell'amore (Venere per i Romani), nata qui, secondo la leggenda, dalla spuma del mare.
2	L'indipendenza è arrivata il 16 agosto 1960.
3	Ultima tappa, in ordine di tempo, è stato l'ingresso ufficiale della Repubblica di Cipro nell'Unione europea il 1° maggio 2004.
4	Che precede di quattro anni l'entrata ufficiale nel sistema dell'euro.
5	Nicosia (Lefkosia) è, dall'XI secolo, la capitale dell'isola.
6	Rimane comunque fuori dall'itinerario che proponiamo la piccola regione del nord, occupata dalla Turchia fin dal 1974.
7	I punti forti della visita a Nicosia sono le mura veneziane (XVI secolo) che abbracciano il nucleo medievale, il Museo archeologico e quello bizantino, la cattedrale secentesca di San Giovann
8	La seconda città dell'isola è Limassol, di origine medievale, con la fortezza dove Riccardo Cuor di Leone avrebbe sposato Berengaria di Navarra.
9	.
10	.
11	.
12	In un'isola il mare è naturalmente l'attrazione principale.
13	E quello di Cipro è stupendo.

Figure 3. Number of periods tab after executing Italian text with Test Calculator.

In the "periods" tab we see how the number of periods is divided. It starts with each capital letter and ends with a full stop. Characters consisting of only dots are themselves a period. So in this example we have a text with thirteen (13) periods. The same goes for the suggestions tab. We observe a detailed situation in the tab "Words - Characters". Here we notice that in the first column with A / A we have the set of words, in the second column the words used in the dictionary and next to the corresponding levels A1, A2, B1, B2, C. At these levels we mark YES if the word is in the dictionary and NO if the word is not in the dictionary. Finally, the last column labeled "outside dictionaries" includes words that do not exist in our database.

Words

Table 1. Word and character tab results.

A/A	Word	A1	A2	B1	B2	C1	except dictionaries
0	Cipro	NO	NO	NO	NO	NO	YES
1	è	NO	NO	NO	NO	NO	YES
2	legata	NO	NO	NO	NO	NO	YES
3	alla	NO	NO	NO	NO	NO	YES
4	mitologia	NO	NO	NO	NO	NO	YES
5	e	YES	YES	YES	YES	NO	NO
6	al	NO	NO	NO	NO	NO	YES
7	nome	YES	YES	YES	YES	NO	NO
8	di	YES	YES	YES	YES	NO	NO
9	Afrodite	NO	NO	NO	NO	NO	YES
10	la	YES	YES	YES	YES	NO	NO
11	dea	NO	NO	NO	NO	NO	YES
12	della	NO	NO	NO	NO	NO	YES
13	bellezza	NO	NO	NO	YES	NO	NO
14	e	YES	YES	YES	YES	NO	NO
15	dell	NO	NO	NO	NO	NO	YES
16	amore	YES	YES	YES	YES	NO	NO
17	Venere	NO	NO	NO	NO	NO	YES
18	per	YES	YES	YES	YES	NO	NO
19	i	YES	YES	YES	YES	NO	NO
20	Romani	NO	NO	NO	NO	NO	YES
21	nata	NO	NO	NO	NO	NO	YES
22	qui	YES	YES	YES	YES	NO	NO
23	secondo	YES	YES	YES	YES	YES	NO
24	la	YES	YES	YES	YES	NO	NO
25	leggenda	NO	NO	NO	NO	NO	YES
26	dalla	NO	NO	NO	NO	NO	YES
27	spuma	NO	NO	NO	NO	YES	NO
28	del	NO	NO	NO	NO	NO	YES
29	mare	YES	YES	YES	YES	NO	NO
30	L	NO	NO	NO	NO	NO	YES
31	indipendenza	NO	NO	NO	NO	NO	YES
32	è	NO	NO	NO	NO	NO	YES
33	arrivata	NO	NO	NO	NO	NO	YES
34	il	YES	YES	YES	YES	NO	NO
35	16	NO	NO	NO	NO	NO	YES
36	agosto	YES	YES	YES	YES	NO	NO
37	1960	NO	NO	NO	NO	NO	YES
38	Ultima	NO	NO	NO	NO	NO	YES
39	tappa	NO	NO	NO	NO	NO	YES

40	in	YES	YES	YES	YES	NO	NO
41	ordine	NO	NO	YES	YES	NO	NO
42	di	YES	YES	YES	YES	NO	NO
43	tempo	NO	YES	YES	YES	NO	NO
44	è	NO	NO	NO	NO	NO	YES
45	stato	NO	NO	NO	YES	NO	NO
46	l	NO	NO	NO	NO	NO	YES
47	ingresso	NO	NO	NO	NO	NO	YES
48	ufficiale	NO	NO	NO	NO	NO	YES
49	della	NO	NO	NO	NO	NO	YES
50	Repubblica	NO	NO	NO	YES	NO	NO
51	di	YES	YES	YES	YES	NO	NO
52	Cipro	NO	NO	NO	NO	NO	YES
53	nell	NO	NO	NO	NO	NO	YES
54	Unione	NO	NO	NO	NO	NO	YES
55	europea	NO	NO	NO	NO	NO	YES
56	il	YES	YES	YES	YES	NO	NO
57	1	NO	NO	NO	NO	NO	YES
58	maggio	YES	YES	YES	YES	NO	NO
59	2004	NO	NO	NO	NO	NO	YES
60	Che	YES	YES	YES	YES	NO	NO
61	precede	NO	NO	NO	NO	NO	YES
62	di	YES	YES	YES	YES	NO	NO
63	quattro	NO	NO	NO	NO	NO	YES
64	anni	NO	NO	NO	NO	NO	YES
65	l	NO	NO	NO	NO	NO	YES
66	enTest Calculatora	NO	YES	YES	YES	NO	NO
67	ufficiale	NO	NO	NO	NO	NO	YES
68	nel	NO	NO	NO	NO	NO	YES
69	sistema	NO	NO	NO	YES	NO	NO
70	dell	NO	NO	NO	NO	NO	YES
71	euro	YES	YES	YES	YES	NO	NO
72	Nicosia	NO	NO	NO	NO	NO	YES
73	Lefkosia	NO	NO	NO	NO	NO	YES
74	è	NO	NO	NO	NO	NO	YES
75	dall	NO	NO	NO	NO	NO	YES
76	XI	NO	NO	NO	NO	NO	YES
77	secolo	NO	NO	YES	YES	NO	NO
78	la	YES	YES	YES	YES	NO	NO
79	capitale	NO	NO	NO	YES	NO	NO
80	dell	NO	NO	NO	NO	NO	YES
81	isola	NO	YES	YES	YES	NO	NO
82	Rimane	NO	NO	NO	NO	NO	YES
83	comunque	NO	NO	YES	YES	NO	NO
84	fuori	YES	YES	YES	YES	NO	NO
85	dall	NO	NO	NO	NO	NO	YES
86	itinerario	NO	NO	NO	NO	NO	YES

87	che	YES	YES	YES	YES	NO	NO
88	proponiamo	NO	NO	NO	NO	NO	YES
89	la	YES	YES	YES	YES	NO	NO
90	piccola	NO	NO	NO	NO	NO	YES
91	regione	NO	NO	NO	YES	NO	NO
92	del	NO	NO	NO	NO	NO	YES
93	nord	NO	YES	YES	YES	NO	NO
94	occupata	NO	NO	NO	NO	NO	YES
95	dalla	NO	NO	NO	NO	NO	YES
96	Turchia	NO	NO	NO	NO	NO	YES
97	fin	NO	NO	NO	NO	NO	YES
98	dal	NO	NO	NO	NO	NO	YES
99	1974	NO	NO	NO	NO	NO	YES
100	I	YES	YES	YES	YES	NO	NO
101	punti	NO	NO	NO	NO	NO	YES
102	forti	NO	NO	NO	NO	NO	YES
103	della	NO	NO	NO	NO	NO	YES
104	visita	NO	NO	NO	NO	NO	YES
105	a	YES	YES	YES	YES	NO	NO
106	Nicosia	NO	NO	NO	NO	NO	YES
107	sono	NO	NO	NO	NO	NO	YES
108	le	YES	YES	YES	YES	NO	NO
109	mura	NO	NO	NO	NO	NO	YES
110	veneziane	NO	NO	NO	NO	NO	YES
111	XVI	NO	NO	NO	NO	NO	YES
112	secolo	NO	NO	YES	YES	NO	NO
113	che	YES	YES	YES	YES	NO	NO
114	abbracciano	NO	NO	NO	NO	NO	YES
115	il	YES	YES	YES	YES	NO	NO
116	nucleo	NO	NO	NO	NO	NO	YES
117	medievale	NO	NO	NO	NO	NO	YES
118	il	YES	YES	YES	YES	NO	NO
119	Museo	YES	YES	YES	YES	NO	NO
120	archeologico	NO	NO	NO	NO	NO	YES
121	e	YES	YES	YES	YES	NO	NO
122	quello	YES	YES	YES	YES	NO	NO
123	bizantino	NO	NO	NO	NO	NO	YES
124	la	YES	YES	YES	YES	NO	NO
125	cattedrale	NO	NO	NO	NO	NO	YES
126	secentesca	NO	NO	NO	NO	NO	YES
127	di	YES	YES	YES	YES	NO	NO
128	San	NO	NO	NO	NO	NO	YES
129	Giovanni	NO	NO	NO	NO	NO	YES
130	La	YES	YES	YES	YES	NO	NO
131	seconda	NO	NO	NO	NO	NO	YES
132	città	YES	YES	YES	YES	NO	NO
133	dell	NO	NO	NO	NO	NO	YES

134	isola	NO	YES	YES	YES	NO	NO
135	è	NO	NO	NO	NO	NO	YES
136	Limassol	NO	NO	NO	NO	NO	YES
137	di	YES	YES	YES	YES	NO	NO
138	origine	NO	NO	NO	NO	NO	YES
139	medievale	NO	NO	NO	NO	NO	YES
140	con	YES	YES	YES	YES	NO	NO
141	la	YES	YES	YES	YES	NO	NO
142	fortezza	NO	NO	NO	NO	NO	YES
143	dove	YES	YES	YES	YES	NO	NO
144	Riccardo	NO	NO	NO	NO	NO	YES
145	Cuor	NO	NO	NO	NO	NO	YES
146	di	YES	YES	YES	YES	NO	NO
147	Leone	NO	NO	YES	YES	NO	NO
148	avrebbe	NO	NO	NO	NO	NO	YES
149	sposato	NO	NO	NO	NO	NO	YES
150	Berengaria	NO	NO	NO	NO	NO	YES
151	di	YES	YES	YES	YES	NO	NO
152	Navarra	NO	NO	NO	NO	NO	YES
153	In	YES	YES	YES	YES	NO	NO
154	un	YES	YES	YES	YES	NO	NO
155	isola	NO	YES	YES	YES	NO	NO
156	il	YES	YES	YES	YES	NO	NO
157	mare	YES	YES	YES	YES	NO	NO
158	è	NO	NO	NO	NO	NO	YES
159	naturalmente	NO	NO	NO	NO	NO	YES
160	l	NO	NO	NO	NO	NO	YES
161	attrazione	NO	NO	NO	NO	NO	YES
162	principale	NO	NO	NO	YES	NO	NO
163	E	YES	YES	YES	YES	NO	NO
164	quello	YES	YES	YES	YES	NO	NO
165	di	YES	YES	YES	YES	NO	NO
166	Cipro	NO	NO	NO	NO	NO	YES
167	è	NO	NO	NO	NO	NO	YES
168	stupendo	NO	NO	NO	NO	NO	YES
169	Ma	NO	YES	YES	YES	NO	NO
170	ci	YES	YES	YES	YES	NO	NO
171	sono	NO	NO	NO	NO	NO	YES
172	anche	YES	YES	YES	YES	NO	NO
173	i	YES	YES	YES	YES	NO	NO
174	Monti	NO	NO	NO	NO	NO	YES
175	Troodos	NO	NO	NO	NO	NO	YES
176	la	YES	YES	YES	YES	NO	NO
177	verdeggiante	NO	NO	NO	NO	NO	YES
178	spina	NO	NO	NO	NO	NO	YES
179	dorsale	NO	NO	NO	NO	NO	YES
180	dell	NO	NO	NO	NO	NO	YES

| 181 | isola | NO | YES | YES | YES | NO | NO |
182	dove	YES	YES	YES	YES	NO	NO
183	è	NO	NO	NO	NO	NO	YES
184	possibile	NO	YES	YES	YES	NO	NO
185	compiere	NO	NO	NO	NO	NO	YES
186	passeggiate	NO	NO	NO	NO	NO	YES
187	alla	NO	NO	NO	NO	NO	YES
188	scoperta	NO	NO	NO	NO	NO	YES
189	dei	NO	NO	NO	NO	NO	YES
190	piccoli	NO	NO	NO	NO	NO	YES
191	borghi	NO	NO	NO	NO	NO	YES
192	che	YES	YES	YES	YES	NO	NO
193	vivono	NO	NO	NO	NO	NO	YES
194	ancora	NO	YES	YES	YES	NO	NO
195	secondo	YES	YES	YES	YES	YES	NO
196	iritmi	NO	NO	NO	NO	NO	YES
197	lenti	NO	NO	NO	NO	NO	YES
198	del	NO	NO	NO	NO	NO	YES
199	passato	NO	YES	YES	YES	NO	NO

Due to the large amount of data in this tab, we have singled out only the second sentence of the Italian text. So, in the image below we see that this sentence has eight words with the corresponding letters in each word. For example, it consists of the first word that has one (1) letter and that is L. Then the second word is indipendenza and it consists of twelve (12) letters and so on.

	Κείμενο προς ανάλυση	Περίοδοι	Προτάσεις	Λέξεις - Χαρακτήρες	Διάρθρωση (Tokenization)
	A/A Πρότασης	A/A Λέξης	Γράμματα		
31	2	1	1	L	
32		2	12	indipendenza	
33		3	1	è	
34		4	8	arrivata	
35		5	2	il	
36		6	2	16	
37		7	6	agosto	
38		8	4	1960	

Figure 4. Outline tab results (Tokenization).

Finally, the software enables us to export a very satisfactory graph to the results, which, after processing all these variables, gives us the cognitive level of the text being entered and the level of difficulty. The words that are in the dictionary are represented in blue and the words that are out of the dictionary are represented in red. So in our example the level of the test is B2, just as the original text of the exam was level B2 with the corresponding level of difficulty of the test. In more detail, we observe the following results: The words that exist at level A1 cover 29% of the text.The words that exist at level A2 cover 35%, the words that exist at level B1 cover 37.5% and finally the words that exist at level B2 cover 41%. Thus, in the specific test of the older years which is level B (B1 & B2) the words that exist collectively at level B correspond to 78.5% of the total words that exist in the dictionary. Only 1.5% are the words of level C and in red are the words that are out of the database.
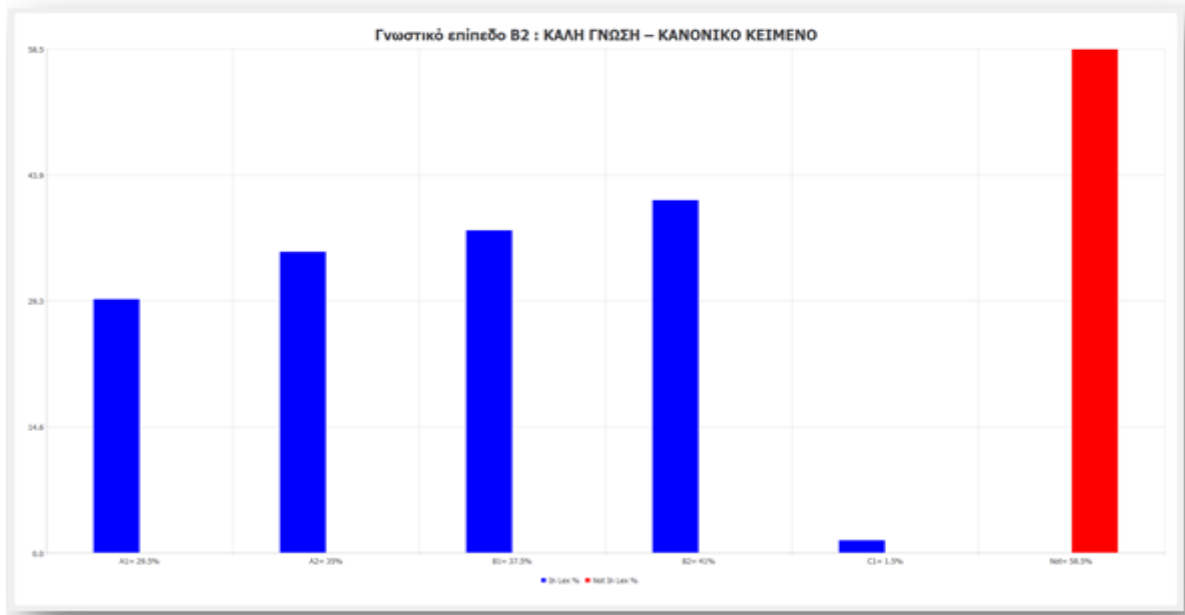
Graph 1. Graph of cognitive level and level of difficulty of the test.

4. Conclusions

The specific software with the parameters we saw above, apart from the analysis of the text that it does, is the only one that measures the degree of difficulty of the tests and the level of difficulty of the readability of the text in Italian. It is easy to use with amazing results. Finally, it contributes positively in the field of education, since it is widely used by the Universities and by the examination committees of the State Certificate of Languages.

References

Bormuth, John R. (2006). Readability: A New Approach. *Reading Research Quarterly*.

Waclawski, E. (2012). How i use it: Survey monkey. *Occupational Medicine*.

Brody, M. J. (2013). Jim Miller & Regina Weinert, Spontaneous spoken language: Syntax and discourse. Oxford: Oxford University Press, 2009. Pp. xiii, 467. Pb. $39.96. *Language in Society*.

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*.

Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*.

DuBay, W. (2004). The Principles of Readability. *Online Submission*.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*.

Marwala, T. (2018a). Gaussian Mixture Models. In *Handbook of Machine Learning*.

Marwala, T. (2018b). Support Vector Machines. In *Handbook of Machine Learning*.

McNamara, D. S., & Graesser, A. C. (2013). Coh-Metrix. In *Applied Natural Language Processing*.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*.

Mikros, G., & Carayannis, G. (2000). Modern Greek corpus taxonomy. *2nd International Conference on Language Resources and Evaluation, LREC 2000*.

Mikros, George, Hatzigeorgiu, N., & Carayannis, G. (2005).Basic quantitative characteristics of the Modern Greek language using the Hellenic National Corpus. *Journal of Quantitative Linguistics*.

Monkey, S. (2014). How does SurveyMonkey adhere to IRB guidelines?

Smith, D. R. (1989). The Lexile Scale in Theory and Practice. Final Report. In *Unknown*.

Waclawski, E. (2012). How i use it: Survey monkey. *Occupational Medicine*.