Modeling the Process of Analysis of Statistical Characteristics of Student Digital Text

Tetiana Shestakevych

Lviv Polytechnic National University, 12, S. Bandert Street, Lviv, 79000, Ukraine

Abstract

The study of student digital texts, generated during distance learning is an urgent scientific and practical task, which involves specialists in various fields, i.e. linguistics, psychology, computer science, etc. The development of distance learning management systems provides researchers and teachers with a tool that simplifies such study. For students of related specialties, participation in the study of their digital texts allows witnessing the diversity of applied study, organized at the intersection of different disciplines. Statistical analysis of the text in combination with the psychological assessment of the speaker identifies ways to find relationships between the relative performance of digital text and the characteristics of the student. It is expedient to use the found dependencies for the personification of pedagogical work with students.

Keywords [1]

Statistical text analysis, text mining, multidisciplinarity, applied linguistics, thematic apperceptive test, TAT, digital text

1. Introduction

The digital text will be text in the digital format produced by students, such as answers to open test questions, essays, and so on. As the share of digital broadcasting has increased significantly since 2020, digital text production has become a common activity. Distance learning is no longer an option but a must. The process of researching digital texts has a powerful set of objects of analysis, and the volume of digital texts is constantly growing. The use of information technology for digital text analysis is an urgent task, in the results of such analysis are interested in economics [1, 2], fashion industry [3] psychologists [4], linguists [5], physicians [6], sociologists [7]. Educators can also benefit from the analysis of digital text produced by students. And not just to determine the level of borrowing in the text or to establish the authorship of the work. Additional benefits of digital text analysis can be obtained in collaboration with experts in other fields. This study will consider approaches to improving the learning process of students based on the analysis of data extracted from digital text, using data analysis methods involving, inter alia, elements of psychological diagnosis.

2. State of Art

The study of texts is an urgent task for the needs of various related fields. The relationship between foreign investment and economic growth was investigated using a text mining approach in [8]. Well-being was investigated via text mining of literature connected with food security [9]. Public opinion on COVID-19 was investigated using text mining in sentiment analysis [7]. The text-based opinion mining technique was used to evaluate the movie and TV show reputation [10], and airlines - based on online reviews [11]. The topic mining analysis, conducted to evaluate the impact of Social Project fostering,

COLINS-2021: 5th International Conference on Computational Linguistics and Intelligent Systems, April 22–23, 2021, Kharkiv, Ukraine
EMAIL: Tetiana.v.shestakevych@lpnu.ua (T. Shestakevych)
ORCID: 0000-0002-4898-6927 (T. Shestakevych)

was conducted by authors in [12]. Trends forecasting based on news mining was used to predict financial changes in Brazilian Market [13]. In [14], the authors investigated the change in time of relevance of research areas in operations research. To effectively collect and classify information, in [15] authors suggested a method based on text mining. An approach to researches in healthcare management and mental health using text mining was suggested in [4]. The analysis of texts written by students was conducted by researchers [16-18].

A study to establish a statistical portrait of a student's digital text is an interesting and urgent task, in particular, because of the possibility of further application of study results not only in linguistics and computer science but also in pedagogy. Moreover, the organization and conduct of such a multidisciplinary study will enable students to trace the relationships between the disciplines they study, give an understanding of the possibilities of applying their knowledge, stimulate scientific activity, encourage compliance with the requirements of academic integrity. The scheme of the study conducted in this work is given in fig. 1.

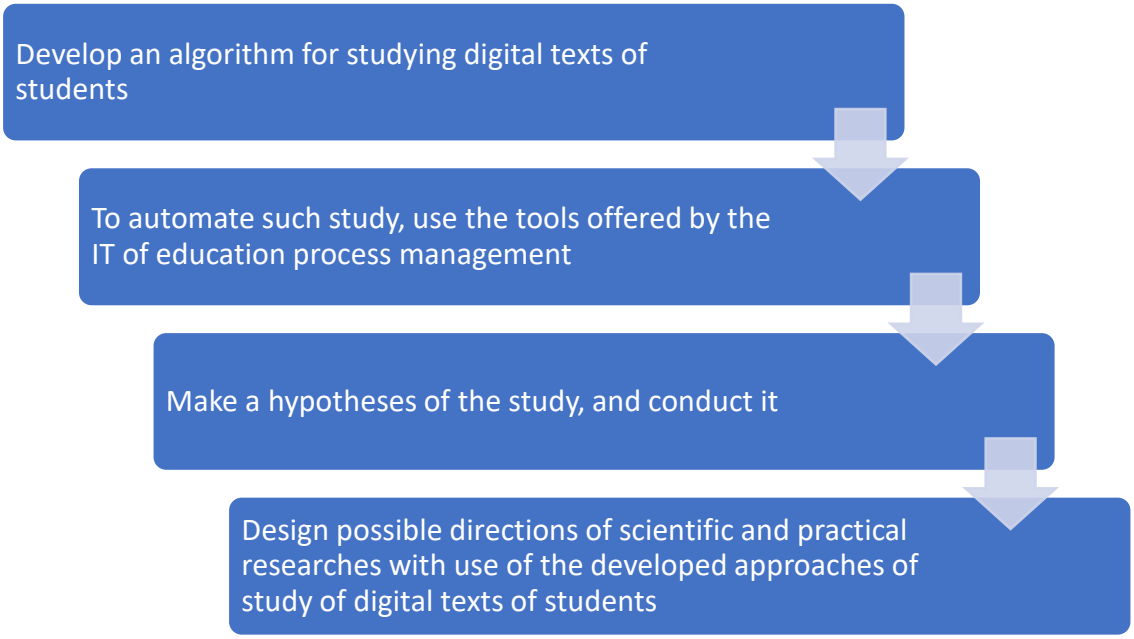

Figure 1: The study scheme

The received texts will be pre-processed, then evaluated statistically by calculating absolute and relative indicators. As such statistical characteristics of the text, one can choose the number of words and sentences in the text, the average length of the sentence, the number of main parts of speech, etc. [19]. Based on these and other statistical characteristics, the coefficients that characterize such text, i.e., the coefficient of diversity, aggression [20, 21], etc. will be calculated.

The study hypotheses testing will be performed using methods and tools of data analysis. Based on the results of such analysis, a decision is made to refute or confirm the hypotheses. The researcher can also assess the shortcomings that arose in the process of such a study to avoid them in the future or to modify the research algorithm.

3. Algorithm for digital text analysis

Stages of study of the digital text of students are presented by such an algorithm.

Algorithm for digital text analysis

Step 1. Formulation of the study purpose, formation of study hypotheses.

Step 2. Selection of stimuli to create a digital text, forming a description of the task.

Step 3. Conducting an independent evaluation of the student.

Step 4. Statistical processing of the text, application of data analysis methods, determination of the characteristics important for the achievement of the purpose of the study.

Step 5. Analysis of study results. Confirmation or refutation of hypotheses. Conclusions on achieving the goal of the study. Conclusions on the process of organizing the study. End of the algorithm.

The described study algorithm was implemented in practice. The study was conducted in February-March 2021, it involved students of the Applied Linguistics department of the National University "Lviv Polytechnic".

4. Description of the study

4.1. The aim of the study. Study hypotheses

The purpose of the study determines the methods and means of achieving it. The use of methods of knowledge extraction, which is based on finding hidden patterns, often makes it possible to ensure the interdisciplinarity of the relevant study. This allows the researcher to go beyond linguistics, to expand the list of problems solved by applied linguistics. We shall state the purpose of the study to verify that digital text reflects the specific characteristics of the speaker. The hypothesis of the study is that there is a relationship between the statistical characteristics of digital text and the psychological characteristics of the speaker.

4.2. Choice of digital stimuli of the study

The work at the hypothesis is based on the use of methods in the field of psychology and will affect the whole process of studying the digital text of students. As an example, consider the features of the use of projective methods of personality assessment, and whether it is possible to find a pattern between the results of such psychological research and the characteristics of digital text. Establishing such a relationship will allow an approximate assessment of the psychological state of the person producing digital text. This can be one of the tools used, for example, when hiring for a job, to monitor the condition of employees, to determine the psychological state of professionals who are often stressed (military, police, teachers, etc.).

Projective methods of psychological research are based on the peculiarities of the functioning of the human brain when a person in the absence of information tends to interpret phenomena or objects based on one`s experience and current psychological state. The person gives meaning to ambiguities, projecting internal conflicts, hidden emotions, and so on. In projective testing, ambiguous drawings, images, and sometimes optical illusions are often the stimulus for expressing opinions (Figure 2-3).

Figure 2: Figure as a stimulus for psychological research

Figure 3: Optical illusion Saxophonist or female face (Image source nesca-newton.com, 2021)

Works [22, 23] are devoted to pojective research methods. One of the methods of projective testing is the Thematic apperception test [24] (TAT), developed in the 1930s by scientists at Harvard University. Such testing is conducted under the guidance of a psychologist, who shows the participant several black and white drawings (one of them is in Fig. 2), these drawings show people (human figures) in casual situations. The research aims to identify the participant with the images, so such images are selected separately for men and women, children, people prone to suicide or depression, and so on. The research participant must invent the story illustrated in the corresponding figure. The answers are also audio recorded to preserve intonation and pauses in speech. On the basis of such research, the psychologist carries out the symptomatic and syndromological conclusion about a condition of the research participant.

To formalize the processes of TAT, it is convenient to use the Petri net (PN1, Fig. 4., Table 1, Table 2). This mathematical abstraction has proven its ability to conveniently visualize the sequence and parallelism of the tasks of a particular process [25].

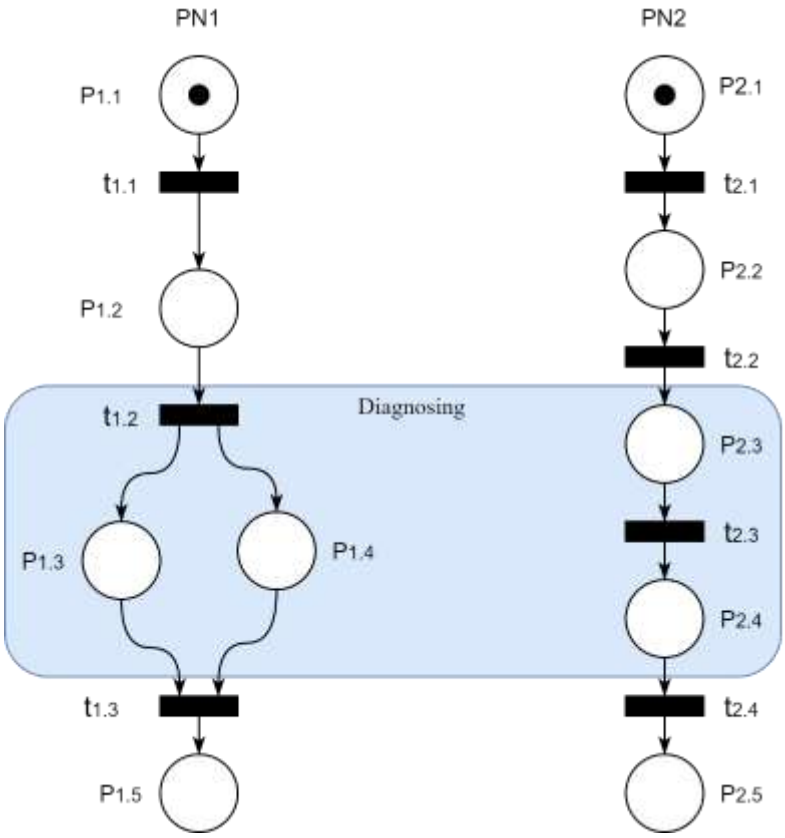

Figure 4: Comparison of the TAT process model, and its modification

Table 1
Transitions in a models TAT and a process of student`s digital text evaluation

Transitions		Meaning of transitions
PN1	PN2	
$t_{1.1}$	$t_{2.1}$	Start a diagnostics
$t_{1.2}$	$t_{2.2}$	A person performs psychological task
	$t_{2.3}$	A person undergoes external evaluation
$t_{1.3}$	$t_{2.4}$	Making complex evaluation

The Petri net $PN1=(P_1, T_1, I_1, O_1)$ models the process of decision making in choosing additional functions, where the set of positions $P_1=\{p_{1.1}, p_{1.2}, p_{1.3}, p_{1.4}, p_{1.5}\}$, the set of transitions $T_1=\{t_{1.1}, t_{1.2}, t_{1.3}\}$; initial marking μ_0 is one chip in position $p_{1.1}$.

Positions in the given Petri net can be interpreted as a condition of event occurrence (Table 2).

Table 2
Positions in a models TAT and a process of student`s digital text evaluation

Positions		Meaning of position
PN1	PN2	
$p_{1.1}$	$p_{2.1}$	Openness for diagnostics
$p_{1.2}$	$p_{2.2}$	A visual stimuli (image)
$p_{1.3}$		Spoken description of an image
$p_{1.4}$		Psychologist's evaluations
	$p_{2.3}$	Digital text
	$p_{2.4}$	Person`s external evaluation
$p_{1.5}$		Diagnosis
	$p_{2.5}$	A pair digital text – person`s external evaluation

From Figure 4 it is convenient to see that the process of drawings describing, it`s audio recording, the participant's cooperation with the psychologist occur simultaneously. One of the options for improving the process of TAT is to automate the work of a psychologist [26-29]. It is not about the use of remote communication technologies, but about changing the focus in the process of forming a description of the image from speech to text. In this way, the psychologist will be able to obtain an additional source of data for analysis. However, this will also mean the need for linguists to process the accumulated data, and this is a task for the psycholinguist. Such a specialist, using psychological and linguistic research methods, will be able to establish the relationship between the statistical features of speech and the psychological state of the patient. Researches [30, 31] were aimed at modeling the workplace of a psycholinguist, taking into account the peculiarities of such an interdisciplinary profession.

The first approach to the implementation of the process of establishing mutual correspondences between the psychological characteristics of the speaker and the statistical characteristics of the digital text produced by him, which is based on thematic apperceptive testing, is to create an information model of the process. In this model, we reflect our vision of the process of accumulating digital text from available sources and using available information technology. Since the results of the study should be used in education, it is logical to use the appropriate information technology to support learning. This can be arbitrary IT, which allows you to collect digital texts from participants in the learning process, as well as record the objective characteristics of the research participant - his age, gender, specialty in which the student is studying, and so on. Lviv Polytechnic National University uses the *Virtual Learning Environment* system on the Moodle platform, and its functionality fully meets the needs of this study. The teacher has the opportunity to create a task within the course, in such a task the corresponding picture is loaded and the field for entering the answer online is provided. The student who completes the task enters a text description and confirms the completion of the task. It is obvious that the functional capabilities of the *Virtual Learning Environment* affect the changes in the TAT process in terms of the accumulation of text examples. It seems logical to model a corresponding modified process by the same means as the TAT model, emphasizing the common and difference of such processes (PN2, Fig. 4, Table 1, Table 2). The Petri net $PN2=(P_2, T_2, I_2, O_2)$ models the process of decision making in choosing additional functions, where the set of positions $P_2=\{p_{2.1}, p_{2.2}, p_{2.3}, p_{2.4}, p_{2.5}\}$, the set of transitions $T_2=\{t_{2.1}, t_{2.2}, t_{2.3}, t_{2.4}, t_{2.5}\}$; initial marking μ_0 is one chip in position $p_{2.1}$.

4.3. Forming a description of the task how to form a digital text

According to the requirements of TAT, the task to describe the figure is presented as simply as possible, without going into explanations. In this study, the task was formulated as follows (translated from Ukrainian):

```
  Write a story based on a picture. The story should
contain:

  What led to this situation?
  What is happening right now?
  What the participants feel and think?
  What will be the consequence of this situation?

  Everything you write is true.
  Time - up to 10 minutes.
```

4.4. Conducting an independent evaluation of the speaker

Students of Lviv Polytechnic National University were involved in the study. The basic characteristics of the study participants are presented in Table 3.

Table 3
The basic characteristics of the study participants

Characteristics of the study	Value
Specialty	Applied linguistics
Academic degree	Bachelor
Semester	8
Participants	62
Female	56
Male	6

In addition to text accumulation, the study of the relationship between the characteristics of the digital text of the speaker and the assessment of the person involves the presence of the assessment. The assessment can be or arbitrary characteristics of the person, or objective (gender, age), or made by an external source (human or automatically calculated, for example, the curator's assessment of student reliability, student educational performance (high/medium/low), etc.). In the case of examining the relationship between the characteristics of the digital text of the speaker and, for example, his psychological state, it involves conducting appropriate psychological research. In the absence of the opportunity to personally assess each student in the current conditions of quarantine, as psychological assessments can be used psychological tests that can be taken remotely, including online tests, if these tests meet the logical requirements:

• The authors of the questionnaires are recognized experts or the questionnaires published in relevant sources of information, i.e. textbooks or manuals that have passed professional testing. The questionnaire has a key.

• In the case of online testing, the reliability of the author of the development and the site where the test is located must be unquestionable. The site should provide the ability to view information about the author of such testing, as well as the availability of feedback.

In this study, *16 individuals* online test was selected for such psychological testing. This online test was developed by NERIS Analytics Limited, which has been developing personal development methods and tools since 2013. The company is registered in Cambridge, UK, and has a LinkenIn profile (https://www.linkedin.com/company/neris-analytics-limited/) and a website (https://www.16personalities.com/uk). This online test (Fig. 5) is adapted to more than 35 languages, including Ukrainian, the company provides access to a number of other tests (paid and free).

Using the tools of the *Virtual Learning Environment*, students receive a link to the appropriate site, and in the comments to the task indicate the results of the test, i.e. one of the sixteen individuals that are formed in four groups (Table 4).

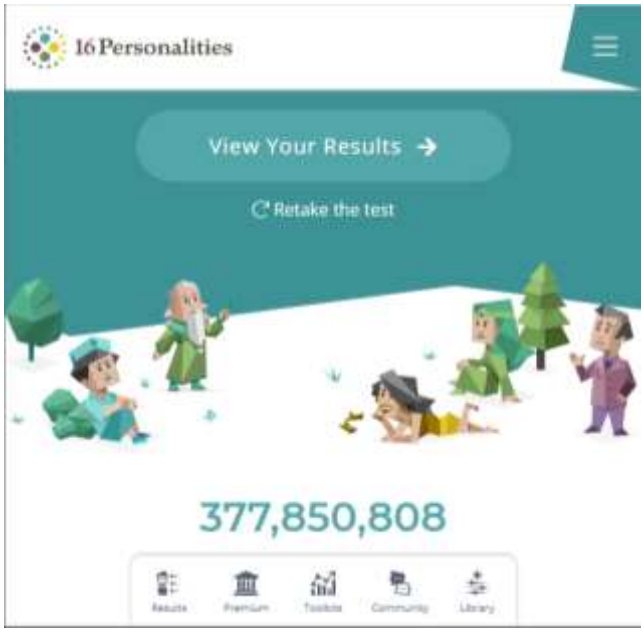

Figure 5: The website of the online testing

The capabilities of the "Virtual Learning Environment" allow the teacher to see the main characteristics of students - participants in the study - the specialty in which students study, the course of study, gender. These characteristics can also be considered as independent assessments of the study participant.

4.5. Accumulation of results of task performance

Students were given 20 minutes to complete the task and take the personality type test. The implementation of the study by means of the *Virtual Learning Environment* does not involve significant time loss either for the formation of the task or for its implementation. Fulfillment of such a task by students now, in the spring of 2021 in such an online environment will not cause concern due to the inability to use the environment, because for over a year it is the basic means of supporting the educational processes of the Lviv Polytechnic National University.

Formal characteristics of the study are given in Table 4.

Table 4
Results of psychological testing 16 personalities

Group of personality types	Number of students (% of the total number of students who took the test)
Analysts (Architects, Logicians, Commanders, Debaters)	6 (12)
Diplomats (Advocates, Mediators, Protagonists, Campaigners)	22 (42)
Sentinels (Logisticians, Defenders, Executives, Consuls)	16 (31)
Explorers (Virtuosos, Adventurers, Entrepreneurs, Entertainers)	8 (15)

The texts of those students who indicated the results of psychological research were admitted to further analysis

4.6. Statistical processing of the text

Statistical evaluation of the text can be carried out according to the following absolute characteristics:

- Number of words in the text
- Number of unique words
- Number of nouns
- Number of verbs
- Number of adjectives
- Number of function words
- Number of pronouns
- Number of sentences
- Number of words with frequency 1;
- Number of words with frequence 10 and more

Based on these characteristics, the following relative indicators can be calculated [20, 21, 32, 33] (Table 5).

Table5

Relative indexes of the text

Index	Formula
Emotionality	$\frac{number\ of\ adjectives}{doubled\ number\ of\ words\ in\ the\ text}$
Emotional stability	$\frac{number\ of\ verbs}{number\ of\ adjectives}$
Level of socialization	$\frac{number\ of\ verbs}{number\ of\ nouns}$
Aggressiveness	$\frac{number\ of\ verbs}{number\ of\ words\ in\ the\ text}$
Vocabulary diversity	$\frac{number\ of\ different\ words}{number\ of\ words\ in\ the\ text}$
Average word repetition	$\frac{number\ words\ in\ the\ text}{number\ of\ different\ words}$
Coefficient of text exclusivity	$\frac{number\ of\ different\ words\ with\ freq.1}{number\ of\ words\ in\ the\ text}$
Text concentration	$\frac{number\ of\ different\ words\ with\ freq.\geq 10}{number\ of\ words\ in\ the\ text}$
Epithetization	$\frac{number\ of\ nouns}{number\ of\ adjectives}$
Verbal definitions	$\frac{number\ of\ adverbs}{number\ of\ verbs}$
Degree of nominality	$\frac{number\ of\ nouns}{number\ of\ verbs}$
Logical coherence	$\frac{number\ of\ function\ words}{number\ of\ sentences}$
Index of intelligibility (readability)	$\frac{average\ number\ of\ words\ in\ sentence}{average\ number\ o\ more\ than\ 6-letter\ words\ in\ a\ sentence}$

Some of these absolute indicators, as well as the indicator of logical coherence, were calculated for four sets (groups) of texts by personality type. Frequency dictionaries were constructed for each of the four sets of texts. This was done using the AntConc software (https://www.laurenceanthony.net/), which allows you to build frequency dictionaries, among other things, texts in Cyrillic (Fig. 6).

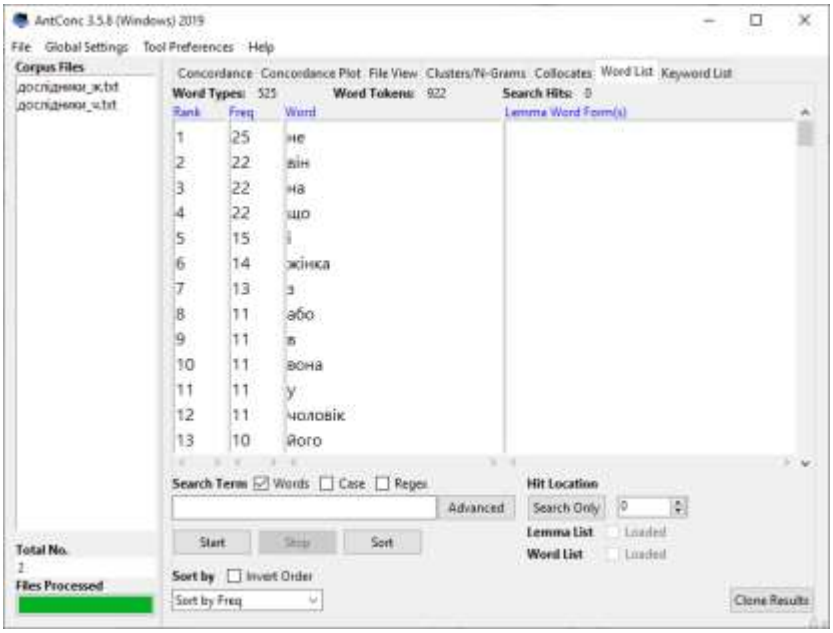

Figure 6: the AntConc software window

Frequency characteristics of the texts are in Table 6 (words with a frequency of more than 3% are given in table 7). The total number of unique words is calculated not as the sum of unique words in each of the types of texts, but on the materials of all texts simultaneously.

Table 6

Frequency characteristics of four groups of texts

Statistical characteristics	Personality type				Total
	Analysts	Sentinels	Diplomats	Explorers	
Total words (% of total)	738 (14)	1193 (23)	2396 (46)	922 (18)	5249 (100)
Unique words (% of total)	440 (60)	619 (52)	1032 (43)	525 (57)	1848 (35)
Top 3% words by frequency (% of total)	136 (13)	329 (24)	681 (51)	248 (12)	1470 (100)

Visually relative indicators from Table 6 is given in fig. 7 and fig. 8.

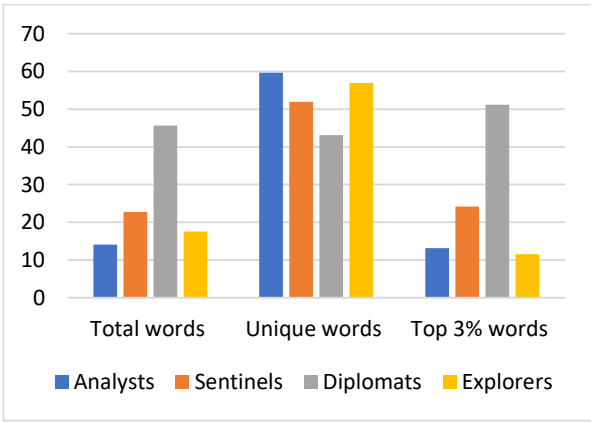

Figure 7: The ratio of statistical indicators in the texts (by personality types)

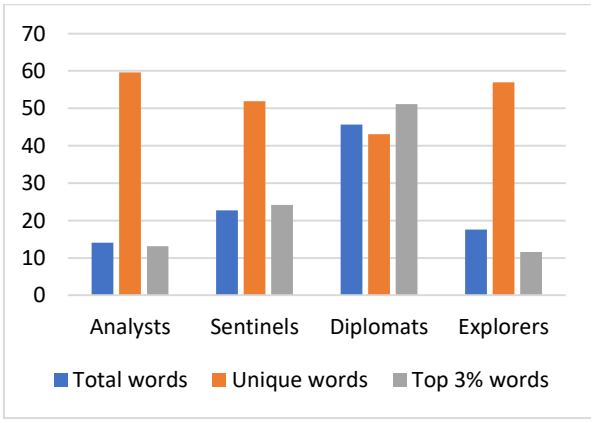

Figure 8: The ratio of statistical indicators in the texts by personality types (by the characteristics of the texts)

Relative indicators of word frequencies of texts (words that make 3 and more percent of the text are chosen).

Table 7

Statistical indicators of the studied texts

	Personality type							
	Analysts		Sentinels		Diplomats		Explorers	
№	Word	Rel. freq. in the text, %	Word	Rel. freq. in the text, %	Word	Rel. freq. in the text, %	Word	Rel. freq. in the text, %
1	на	5	що	5	і	7	не	5
			чоло					
2	що	5	вік	5	на	5	він	4
3	не	5	на	4	що	5	на	4
4	і	4	та	4	не	4	що	4
	жінк				чолов			
5	а	3	і	4	ік	4	і	3
	чоло							
6	вік	3	не	4	жінка	4	жінка	3
7	у	3	його	3	з	4		
8	з	3	він	3	в	3		
9			в	3	та	3		
			жінк					
10			а	3	він	3		
11			з	3	у	3		
12					вона	3		
13					його	3		

Words with a singing frequency of more than 3%, which were in each of the types of texts, are a *woman (*жінка*), and (*і*), on (*на*), not (*не*), what (*що*).*

The number of nouns, pronouns, and function words in the studied texts is given in Table. 8.

Table 8

Number of nouns, pronouns and service words

	Personality type							
	Analysts		Sentinels		Diplomats		Explorers	
	Word	Rel. freq. in the text, %	Word	Rel. freq. in the text, %	Word	Rel. freq. in the text, %	Word	Rel. freq. in the text, %
Nouns	28	21	46	18	83	16	14	12
Pronouns	0	0	41	16	88	17	22	18
Function words	108	79	163	66	358	67	84	70

4.7. Analysis of research results

From the results, we can draw the following conclusions.

Most of the students who participated in the study have a Diplomat personality type (42%), these students are the authors of the largest share of words (46%), and the texts contain the lowest rate of uniqueness (43%). The texts of this group differ significantly in the proportions of the total share of

words among all texts (46%) to the number of unique words (43%), this ratio can be given as 46:43, while for other texts this ratio is 14:60 (Analysts), 23:52 (Sentinels), 18:27 (Researchers).

The fewest students who participated in the study have the personality type Analyst (12%), these students are the authors of the smallest share of words (14%). The texts contain the highest rate of uniqueness (60%).

Proportional for all groups of texts are the indicators of the share of the number of words and the number of words with a frequency of more than 3%. The inverse proportion is observed for the dependence of the share of words in total - the number of unique words.

There were no verbs or adjectives among the words with a frequency of more than 3%, which was true for all types of texts. Comparing the number of nouns, pronouns, and the function words, we note for fairness approximately the same ratio of the number of nouns and pronouns was not performed for texts such as personality Analysts. In such texts, there is the highest share of nouns and the lowest of pronouns. The total share of words per pair of nouns-pronouns is the lowest for texts of the Analysts personality type (21%), for other types of texts such a total share is 30-34%.

Thus, we can say that the research hypothesis is confirmed and there is a relationship between the statistical characteristics of digital text and the psychological characteristics of the speaker. Of course, it is necessary to take into account the volume of research material, the conditions of the study, the restrictions imposed on the research material, and so on. However, more important than finding hidden dependencies in digital texts, for this study is the opportunity to show the correctness of the approach to the analysis of the digital text, and options for its IT support.

5. Conclusions

To investigate the relationship between the statistical characteristics of a student's digital text and some independent assessment of the student's personality, an algorithm was proposed, that involves the formation of multiple student-generated texts, the assessment of students' psychological characteristics, and statistical analysis of digital texts.

Among the results of the study is the identification of several characteristics of the texts, which differ significantly for the different types of the speaker personality. For example, the ratio of the total share of words among all texts to the number of unique words, the total share of words per pair noun-pronoun.

Areas of further research include expanding the volume of digital texts, developing a database and data warehouse schema to preserve the evaluation results and statistical parameters of texts, introducing the language corpus into the study, and so on. Current hypotheses of future research include the relationship between the statistical characteristics of digital text in native and foreign languages, the dependence of statistical parameters of the text on the objective indicators of the speaker - age, course, specialty, science, etc., as well as ensuring the diachronicity of relevant research.

6. References

[1] S. Ranfagni, M. Faraoni, L. Zollo, V. Vannucci. “Combining online market research methods for investigating brand alignment: the case of Nespresso.” British Food Journal 123 (13) (2021): 37- 58.

[2] G. Szymanski, P. Lipinski, Model of the effectiveness of Google Adwords advertising activities, in: Proceedings of the 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2018, Lviv, Ukraine, 2018, vol. 2, art. no. 8526633, pp. 98-101.

[3] P. Casadei, N. Lee. “Global cities, creative industries and their representation on social media: A micro-data analysis of Twitter data on the fashion industry.” Environment and Planning A 52 (6) (2020): 1195-1220.

[4] C. C. Ghosh, D. McVicar, G. Davidson, C. Shannon. “Measuring diagnostic heterogeneity using text-mining of the lived experiences of patients.” BMC Psychiatry, 21 (1), art. no. 60 (2021).

[5] L. Radková, M. Místecký. “Words of warfare: Life of mission soldiers in the perspective of quantitative linguistics.” Bohemistyka (1) (2021): 7-26.

[6] S. Mijomanović, I. Aleksić-Hajduković, D. Sinadinović. “CASE REPORTS IN DENTAL MEDICINE: A GENRE ANALYSIS” ESP Today 9 (1) (2021): 129-149.
[7] S. Albahli, A. Algsham, S. Aeraj, M. Alsaeed, M. Alrashed, H.T. Rauf, M. Arif, M. A. Mohammed. “COVID-19 Public Sentiment Insights: A Text Mining Approach to the Gulf Countries.” Computers, Materials and Continua 67 (2), (2021): 1613-1627.
[8] C. Cicea, C. Marinescu. “Bibliometric analysis of foreign direct investment and economic growth relationship. A research agenda.” Journal of Business Economics and Management 22 (2) (2021): 445-466.
[9] M. W. Cooper, M.E. Brown, M. T. Niles, M. M. ElQadi. “Text mining the food security literature reveals substantial spatial bias and thematic broadening over time.” Global Food Security, 26, art. no. 100392 (2020).
[10] A. Benlahbib, E. H. Nfaoui. “MTVRep: A movie and TV show reputation system based on fine-grained sentiment and semantic analysis.” International Journal of Electrical and Computer Engineering 11 (2), 2021: 1613-1626.
[11] H.-J. Kwon, H.-J. Ban, J.-K. Jun, H.-S. Kim “Topic modeling and sentiment analysis of online review for airlines.” Information (Switzerland) 12 (2), art. no. 78 (2021): 1-14.
[12] M. P. Bach, C. M. Pulido, D. S. Vugec, V. Ionescu, G. Redondo-Sama, L. Ruiz-Eugenio. “Fostering social project impact with Twitter: Current usage and perspectives.” Sustainability (Switzerland) 12 (15), art. no. 6290 (2020).
[13] J. J. Duarte, S. Montenegro González, J. C. Cruz, Jr. “Predicting Stock Price Falls Using News Data: Evidence from the Brazilian Market.” Computational Economics 57 (1) (2021): 311-340.
[14] R. Romero-Silva, S. de Leeuw. “Learning from the past to shape the future: A comprehensive text mining analysis of OR/MS reviews.” Omega (United Kingdom) 100, art. no. 102388 (2021).
[15] Y. Lee, J. Cho. “Web document classification using topic modeling based document ranking.” International Journal of Electrical and Computer Engineering 11 (3) (2021): 2386-2392.
[16] A. Petchprasert. “Utilizing an automated tool analysis to evaluate EFL students’ writing performances” Asian-Pacific Journal of Second and Foreign Language Education 6 (1), art. no. 1 (2021).
[17] E. M. Iñesta Mena, L. I. Montes. “Stated practices on the teaching and assessment of oral communication in primary education [Prácticas declaradas sobre enseñanza y evaluación de la oralidad en educación primaria].”, Tejuelo, 33 (2021): 281-318.
[18] S. Maksymenko, O. Nemesh, K. Maksymenko, A. Nabochuk. “Psychological peculiarities of understanding by students internet texts.” Psycholinguistics 28 (1) (2020): 187-213.
[19] S. Buk, Y. Krynytskyi, A. Rovenchak. “Properties of autosemantic word networks in Ukrainian texts.” Advances in Complex Systems, 22 (6), art. no. 1950016, (2019).
[20] A. Rovenchak, S. Buk. “Part-of-speech sequences in literary text: Evidence from Ukrainian.” Journal of Quantitative Linguistics 25 (1) (2018): 1-21.
[21] S. Buk, A. Rovenchak. “Statistical parameters of Ivan Franko's novel Perekhresni stežky (The Cross-Paths).” Exact Methods in the Study of Language and Text: Dedicated to Gabriel Altmann on the Occasion of his 75th Birthday (2011): 39-48.
[22] H. Jin, Z. Li, W. Liu. “An Value Study of Tree-drawing Test in the Screening of Depression among the Elderly in Communities.” Chinese General Practice 23 (32) (2020): 4034-4038.
[23] S. A. Nita, R. I. Teleanu, O. A. Bajenaru. “Psychological evaluation in childhood migraine – Steps towards a psychological profile of the child migraineur.” Romanian Journal of Neurology/ Revista Romana de Neurologie 19 (1) (2020): 21-26.
[24] S. R. Jenkins. “The Narrative Arc of TATs: Introduction to the JPA Special Section on Thematic Apperceptive Techniques.” Journal of Personality Assessment 99 (3) (2017): 225-237.
[25] T. Shestakevych, V. Pasichnyk. “The use of Petri Nets for inclusive education IT-support.” ECONTECHMOD, Vol. 04, No. 2 (2015): 33–38.
[26] V. Vasyliuk, Y. Shyika, T. Shestakevych, Information system of psycholinguistic text analysis, in: CEUR Workshop Proceedings, 2604, 2020, pp. 178-188.
[27] K. Fornalczyk, P .Napieralski, D. Szajerman, A. Wojciechowski, P. Sztoch, J. Wawrzyniak, Stereoscopic image perception quality factors, Proceedings of the 22nd International Conference on Mixed Design of Integrated Circuits and Systems, MIXDES 2015, 2015, art. no. 7208495, pp. 129-133.

[28] K. Fornalczyk, A. Wojciechowski, Robust face model based approach to head pose estimation, Proceedings of the 2017 Federated Conference on Computer Science and Information Systems, FedCSIS 2017, 2017, art. no. 8104720, pp. 1291-1295.
[29] D. Szajerman, P. Napieralski, J.-P. Lecointe. "Joint analysis of simultaneous EEG and eye tracking data for video images." COMPEL - The International Journal for Computation and Mathematics in Electrical and Electronic Engineering, 37 (5) (2018): 1870-1884.
[30] V. Vasyliuk, Y. Shyika, T. Shestakevych, Modelling of the Automated Workplace of the Psycholinguist, in: Proceedings of the 2020 IEEE 15th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2020, Lviv, Ukraine, 2020, vol. 1, art. no. 9321956, pp. 276-279.
[31] D. Szajerman, M. Warycha, A. Antonik, A. Wojciechowski. "Popular brain computer interfaces for game mechanics control." Advances in Intelligent Systems and Computing, 506, (2017): 123-134.
[32] I. Kamińska-Szamaj, Różnice leksykalne między stylami funkcjonalnymi polszczyzny pisanej: analiza statystyczna na materiale słownika frekwencyjnego, Wrocław, 1990.
[33] M. Ruszkowski. "The Category of Generation in Polish Linguistic Studies (Current Status and Perspectives)." Respectus Philologicus 27 (2015): 147-155.