THE AUTHOR'S QUANTIFIED APPROACH – VIA THE SIMILITUDE REPORT: FROM REFERENCES TO CREATIVITY

Ţaicu Marian [1], Gheorghe Săvoiu [2]

[1,2]University of Pitesti, Faculty of Economic Sciences and Law, 1st Targul din Vale, Piteşti, Argeş, Romania,
[1]e-mail: taicumarian@yahoo.com; [2]e-mail: gsavoiu@yahoo.com

Abstract. *The authors analyze the structure of a standard article based on the similarity report made by any prestigious publication before a peer review and eligibility check. After a brief introduction, in which the general structure of the article, in the authors' view, and the final structure of the article, subsequent to the control of publishers or journals, are seen in parallel, the central section of the article presents several similarity reports with the Plagiarism Detector (used in demo version). References and citations, together with their limiting weight in the text of the article, are the main topics for discussion. Some final remarks highlight the relative relevance of using limiting thresholds for the amount of references and citations in publications.*

Keywords*: references, citations, bibliography, plagiarism detector, similitude report, scientific research*

1. INTRODUCTION

An article, once conceived, broadly complies with the sections of a classic research report, prior to the editing and publication, starting with the title and summary together with its specific keywords, plus a literature review section, then another one which is methodological, and especially one devoted to results and discussions, and finally conclusions or remarks are necessary, which are focused on the limits of the paper and the perspectives of the synthesized scientific research. In the final stage, the publishing house or journal analyzes the article from the perspective of the standard structure – namely referenced, plagiarism and original. The structural analysis of the similarities, mirrored by the analysis of classical sections, starts from the information defined by the anti-plagiarism report, made by any prestigious publication, prior to a peer review and a necessary eligibility check.

There is no direct or indirect relationship between the classical sections of the article and its standard structures, but rather the three structures are found in each section, or in other words all the sections are x-rayed in terms of percentages related to the level of structures resulting from the similarity investigation: referenced, plagiarism and original. The present article exemplifies this structure by means of an anti-plagiarism software, which also quantifies the actual amount of references and citations or quotations, generating a useful and original discussion about the optimal references and citations.

## 2. THE RELATIVE OPTIMUM OF REFERENCES AND CITATIONS ACCORDING TO THE SIMILARITY REPORT

Two major features of a carefully and acurately designed article, rigorously drafted and published in a journal that enjoys real prestige, consist of exemplification and argumentation with a probative role, both pragmatically realized with the help of other works (books, articles), which thus support the importance of the themes, topics, hypotheses, methods, models, experiments, simulations of results, or even findings of the research. This practice becomes more convincing and validates the hypotheses and opinions of the author or authors. Such an approach constantly seeks to identify an optimum of references and citations or quotations, often exemplifying in the form of references to other works or citations and quotes from other texts in various books or articles [1]

It should be remembered that if the list of references and the list of citations can coincide, and even do coincide in the modern conceptualization of an article, they nevertheless differ clearly in relation to the bibliography of an article, the former listing, in the end, all the articles quoted in the text of the book or article, while a bibliography covers all the papers (books, articles, etc.) that an author has consulted in the preparation of his/her investigation, whether or not he/she has generated references or quotes from them.

The reference or citation, in the structural context of the similarity report, is quantified based on the existence or not of quotes (i.e. inverted commas) that delimit the actual quotation, since the fact they are missing places the respective text outside the referenced structure, and into the structure of plagiarism. The bibliography loses its major significance, acquired in classical sections, while references and citations or quotations turn into a relevant structure relative to the creativity or originality of the paper.

The Plagiarism Detector software, or software package, provides, in addition to the total and per-section analysis of the article and book text, a quantification of each structure, both in terms of the number of similar or identical words in phrasing (for suspected theft of ideas and detection of area of plagiarism), and in terms of percentage (for all structures from referenced, to plagiarism and original). This can be exemplified in the content of a final report of the Plagiarism Plagiarism Detector software, which is relatively incomplete in its available demo version, being numbered (e.g. Detector v. 1020) and redefined as the Originality Report, as a scanning process for an analyzed document (dated for the year and day in the calendar, e.g. 30.08.2017, but also for the time of the day, e.g. 17:33:39). The report is nominated and numbered successively, in an exemplifiable form such as

Report 1 Name and surname 30.08.2017.docx., being synthesized graphically:
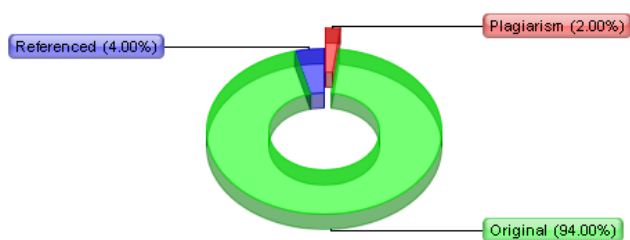
Source: Distribution graph: Comparison Preset: Word-to-Word. Detected language: Romanian
Fig 1. Generated by Plagiarism Detector (Demo Version) (Warning: Demo Version reports are incomplete)

Using this program to check a text involves going through several steps:
Step 1. Add documents;
Step 2. What to detect. The user must choose between Word-to-Word detection and Text Rewrite detection;
Step 3. Select the Check Type
So, the program allows one to choose how to detect plagiarism: either "Word-to-Word", which provides maximum accuracy and is recommended for areas like Sciences, or "Text Rewrite", which offers maximum detection and is recommended for the Arts domain.
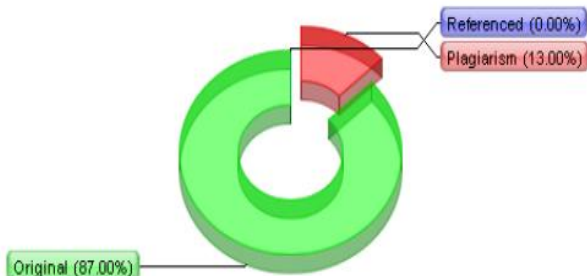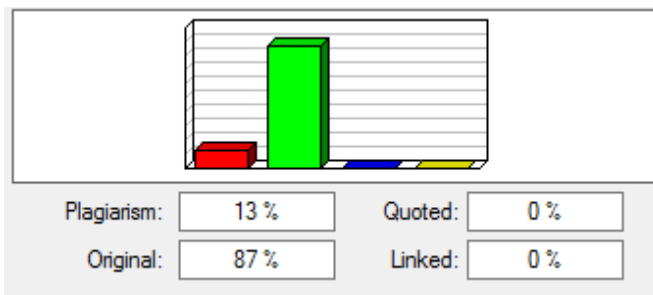

Source: Distribution graph: Comparison Preset: Text Rewrite. Detected language: Romanian
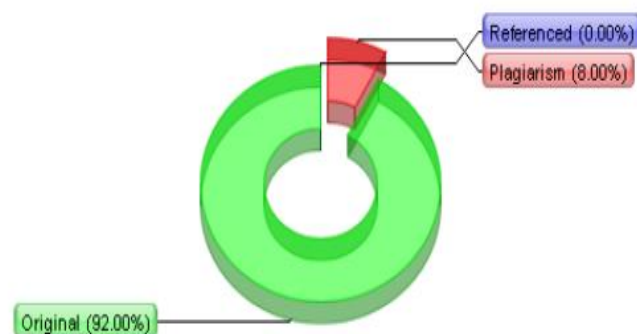Fig 2. Generated by Plagiarism Detector (Demo Version) (Warning: Demo Version reports are incomplete)

In the panel of viewing the reports for the latest verified documents, the weights of plagiarism, citations / quotes, original text, and links are graphically presented as columns, as shown in Figure 3.

Using plagiarism detection options (Word-to-Word or Text Rewrite) can generate slightly different results. The authors used the same text containing plagiarism elements, and checked it using both options. Figure 4 shows the results obtained for the same text as that in Figures 2 and 3, but the option went for Word-to-Word verification. A slight decrease in the percentage indicated for plagiarism can be noted, which is but normal given its multiple forms.

Source: Distribution graph: Comparison Preset: Text Rewrite. Detected language: Romanian
Fig 3. Generated by Plagiarism Detector (Demo Version) (Warning: Demo Version reports are incomplete)

Source: Distribution graph: Comparison Preset: Word-to-Word. Detected language: Romanian
Fig 4. Generated by Plagiarism Detector (Demo Version) (Warning: Demo Version reports are incomplete)

Creativity has a multitude of definitions in the literature, but these can be summed up as the ability to create new ideas, concepts, and methods. Creativity is closely linked to originality, the two concepts lying at the basis of the progress and diversity of human society. Both concepts have been intensely debated in the last centuries, and are viewed from different angles by specialists in various fields.

There is a close connection between creativity and originality: we cannot talk about creativity without originality, or vice versa.

The development of ICT has made it possible to produce collections of documents unprecedented with regard to size, and the field of scientific literature is no exception. Scientific papers can be organized or grouped within these collections based on criteria such as keywords or JEL codes, but in recent years there have been proposals for grouping them based on the analysis of the references of each work, using Artificial Neural Networks (ANNs). [3]

Plagiarism is a detrimental phenomenon in any field, including the world of science. The spread of the Internet has favoured the appearance of various forms of plagiarism which are hard to detect by classical means. That is why plagiarism detection software is now used that offers maximum efficiency in a short time.

The Plagiarism Detector program can detect plagiarism in text, online and offline documents, and can be successfully used by universities, publishers, journals, or anyone interested

in the originality of a document. This program uses the largest digital database available to the public in the world – the one presupposed by search engines.

One standard that can measure its success is the fact that the program is being used by customers in 150 countries.

The Allen Institute of Artificial Intelligence [4] has developed the Semantic Scholar project [5], a non-profit, free of charge academic search engine based on artificial intelligence. This search engine provides access to millions of scientific papers using machine learning techniques, and calculates the number of citations for these papers taking into consideration only quality papers.

5. SOME FINAL REMARKS

If there are two images of a paper or article and, implicitly, as many approaches, which are sharply different, the first being focused on the classical sections and the second on the structures quantified by the similarity report, what finally manages to secure their essential cohesion is references or citations and creativity or originality. In any section, the references or citations along with creativity or originality, as exclusive solutions to plagiarism, ensure the continuity of the methodological, argumentative or motivational approach, synthetically and analytically investigative. Even though the authors of the present article consider that a general optimization of the percentage restriction of structures is not justified, the option of a Pareto (20/80) ratio between the referenced and the original matter seems to be a real solution, which is able to ensure the balance of an article. In addition to Plagiarism Detector, there are many other programs capable of identifying correct citations or quotes and incorrect quotes (which are basically transformed into plagiarism).

The article structures can be multiplied relative to the software package or software used. The first structure identical to the referenced matter is given by the percentage values of the citations / quotes, plagiarism is identified as the structures of identity (or absolute similarity) structures, and the relative similarity with the texts found after the confrontation as potential sources of plagiarism, and the original is relativized and renamed as a structure defined by indecision, which is eventually quantified as the percentage of phrases or sentences that were not found in the contrasting base and/or the Web.

6. References

[1] Dinu, V., Săvoiu, G., Dabija, C (2016). *A concepe, a redacta şi publica un articol ştiinţific. O abordare în contextul cercetării economice*, Bucureşti: Editura ASE.

[2] Plagiarism detector software (2017). [online] Available at:https://plagiarism-detector.com/d2/c/en/plagiarism-detector-download-demo.php [Accessed November 23, 2017].

[3] Magali Rezende Gouvêa Meireles, Beatriz Valadares Cendón (2015). Citation-Based Document Categorization: An Approach Using Artificial Neural Networks, *Qualitative and Quantitative Methods in Libraries (QQML) Special Issue Bibliometrics and Scientometrics:* 71-79

[4] http://allenai.org/index.html

[5] https://www.semanticscholar.org/