

ABS 214

by lcels_2 Abs 214

Submission date: 30-Sep-2020 04:24PM (UTC+0700)

Submission ID: 1401159474

File name: full_paper_abs-214_8956834453.docx (592.15K)

Word count: 3072

Character count: 17682

Computer-Based System Design for Writing Assessment in Indonesian Language

Sri Kusuma Winahyu*, Endry Boeriswati and Fathiaty Murtadho

¹¹
Universitas Negeri Jakarta, Jakarta Timur, Indonesia

*Corresponding author

srikusumawinahyu_9906917033@mhs.unj.ac.id

HP: +6281280335158

ABSTRACT

This paper describes the design of a computer-based writing assessment system in Indonesian language, especially assessment of opinion article writing. The system designed for assessment is in the form of a computer based test (CBT) system with machine rater and human rater. The computational process is carried out by applying preprocessing steps in natural language processing (NLP). The need analysis of the Indonesian language teachers regarding the assessment of writing opinion articles by grade XII of high school students which is used as the basis for this study shows that in this assessment system it is necessary to include trigger questions and the completeness of the assessment components. These two things are used by researchers to plan the design of the assessment instrument. This assessment design is considered important to be initiated considering that so far technology has become an absolute necessity in the teaching and learning system. The benefits of this system when associated with learning Indonesian language are in terms of efficiency and objectivity in teacher assessments of grade XII in accordance with the basic competencies in the 2013 Curriculum to be able to produce opinion articles. This research is also relevant to conditions today, when the Covid-19 virus outbreak occurred, which then had an impact on conditions that “forced” teachers and students to carry out the teaching and learning process online.

Keywords: CBT, NLP, opinion articles, machine rater, and human rater.

9

Introduction

Writing is one of the language skills that students must master in the learning context at school, in addition to other skills, such as listening, reading, and speaking. In normal situations, skills learning activities are carried out in daily life at school face-to-face between students and teachers in the classroom. However, all countries in the world, including Indonesia, are currently facing the covid-19 pandemic, hence all face-to-face learning activities must be replaced by technology based distance learning. This condition obliged of experts and researchers in the field of education to think and realize various technology to facilitate the learning process.

The computer-based writing assessment system is an important technology to use today. This system has actually been initiated for a long time, which began with the emergence of automatic scoring in the 1960s. Then in 1970 came the automatic essay grading machine technology developed by the University of Connecticut. Automatic essay assessment is a measurement technology in which computers evaluate written work (Shermis & Burstein, 2003 in Shermis, 2010). In 1990 the system was revitalized with two technologies, such as microcomputers and networks. Microcomputers allow text to be written electronically and networks allow text to be sent for the assessment process. However, the system development that was initially carried out to be applied in English still has limitations, that it cannot yet grasp the connotative meanings of sentences. The system generally only assesses the mechanical elements and vocabulary.

This limitation is also found in research on computer-based writing assessment that has been developed in Indonesian language. Two of these studies were conducted by Luh Anik (2019) and Anak Agung Putri Ratna (2015). Luh Anik developed a computational Indonesian spelling editing application (SIPEBI). The application checks the mechanical elements and

words formation in the sentences produced by the Indonesian language counseling participants. Meanwhile, Ratna developed a web-based automatic essay grading system (SIMPLE-O). This system assesses student essays using Latent Semantic Analysis (LSA) which emphasizes the words in any text without paying attention to the rhetoric of the sentence and its linguistic characteristics, such as grammar, how it is written, and what the order of the words in a sentence looks like. Assessment in ⁴ this system is carried out by comparing student answers with the best answers from lecturers who have previously been inputted into the system. SIMPLE-O has also been compared with human raters and has an accuracy of above 86%. Even so, the two assessment systems are not sufficient if applied to writing assessments in the context of learning Indonesian in schools because to assess students' writing results in the context of Indonesian language subjects, a more detailed assessment component is needed that is able to clearly show the performance of each student. This is in accordance with the opinion of Lyons (2014), that writing assessment refers to all forms, contexts, methods / strategies that show the quality and characteristics of a person's written text. Writing assessment refers to the field of language study and is part of language skills and abilities. Writing assessment is also included in measuring performance.

The development of computer-based writing skills assessment in Indonesian language in linguistic studies and computational technology has not been widely found. In fact, computer-based writing assessment technology is needed to facilitate teacher, especially during the pandemic due to the current corona virus where the teaching and learning process is carried out using an online system. One of the things that might complicate the development of a complete computational system is the difficulty of machines assessing the results or answers of writing tests in the form of sentences, paragraphs, even whole text or articles so that the assessment involves many components and considerations. Brown (2003) suggests that a possible form assessment is a combination of a machine rater and a human rater with a computer based test

(CBT) system. The use of two raters: machine rater and human rater in a computer-based writing assessment system has been carried out to assess writing proficiency in ¹⁴ the Test of English for International Communication (TOEIC) organized by the Educational Testing Service (ETS) (Everson, 2010). ETS uses an ⁷ online scoring network (OSN) system. The answers to the writing test from the participants were sent via the internet, then the human rater assessed the scoring interface. Human assessors in the TOEIC writing assessment process who are separate from the test takers, even do not know each other, are always trained and calibrated regularly.

This article describes a qualitative study that aims to explore the teacher's needs for a computational writing assessment system and design an Indonesian writing assessment system according to teacher needs. The type of text as the object of this research is an opinion article produced by grade XII of high school students, according to the 2013 curriculum for Indonesian language lessons. This study also ¹⁰ seeks to answer two questions: 1) what are the constraints faced by the teacher while assessing student opinion articles manually and 2) how to design an assessment system for writing opinion articles in Indonesian based on the constraints faced by the teacher.

Material and Methods

In this study, researchers used a qualitative descriptive method. Qualitative research is an exploratory study aimed at understanding and exploring certain subjects and/or objects (Creswell, 2018). In this case the researcher tried to explore the needs of Indonesian language teachers for a computer-based writing assessment system and its components based on their experience when assessing student opinion articles manually. That experience is reality as it appears to the teacher. This is a phenomenological category in qualitative methods (Borg and Gall, 2002). To explore the teacher's need for the importance of developing computer based

writing assessment tools, the researcher applies participants' meaning, where the researcher focuses on studying the needs of participants or Indonesian language teachers. As a characteristic of qualitative research that starts from empirical facts, researchers collect data from the field, study, analyze, interpret it, then draw conclusions from phenomena in the field. The data analysis was carried out to find the meaning which was the result of the research and used for program development.

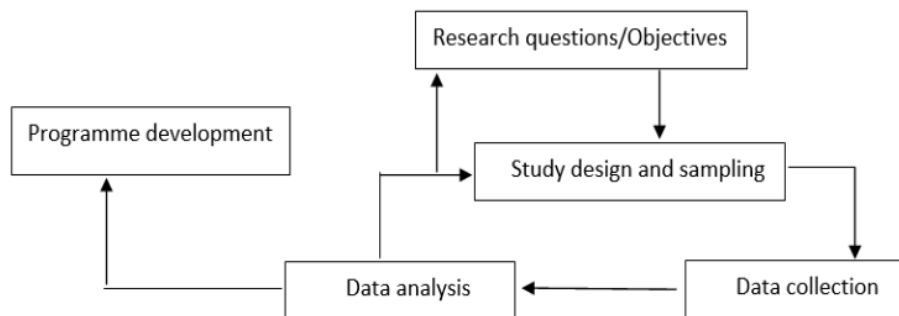


Figure 1. Qualitative research process

From Figure 1 it can be explained that this qualitative research process begins with determining the research question and objectives. The research questions in this research are what are the obstacles faced by the teacher while assessing student opinion articles manually and how to design a computer-based scoring system for writing Indonesian opinion articles based on the constraints faced by the ¹teacher. The purpose of this study is to explore the teacher's needs for a computational writing assessment system and to design an Indonesian writing assessment system according to teacher needs.

²²In order to answer the research questions and achieve the above objectives, the researcher designed the study and assigned the sample. The research that was designed was a qualitative descriptive study by exploring information on needs analysis and the final result in the form of a program design based on needs analysis, while the sample set was a grade XII Indonesian

teacher. Furthermore, the researcher collected data through literature study and interviews. A literature study was conducted regarding the basic competencies of grade XII of high school students and interviews were conducted to explore the needs of Indonesian language teachers for computer-based writing assessment tools. The results of the interviews are sorted and grouped with one another which has similarities. In data analysis, the researcher explores one by one the answers from the teachers who have been grouped and relates them to the components that must be built in the assessment system. The results of this data analysis are then transformed into a computer-based writing assessment program development design.

Results and Discussion

Literature study, which is one of the steps for exploring data in this research, yields information that in the ³ Regulation of the Minister of Education and Culture of the Republic of Indonesia Number 24 of 2016 there are basic competencies for grade XII of high school in the 2013 Curriculum, that ¹ students are required to be able to compose opinions in the form of articles such as in magazines. Meanwhile, the results of grouping the teacher's answers in the interview are that during the manual assessment of the results of writing student opinion articles, the teacher often experiences problems: 1) student statements or thesis are not well structured if the teacher does not provide trigger questions; 2) teachers do not always use the assessment components completely; 3) teachers often find it difficult to complete assessments quickly; and 4) teachers sometimes find it difficult to be objective. These constraints are then processed in data analysis and become forms of teacher needs for components in the computer-based writing assessment system so that in designing a computer-based opinion article writing assessment system, the researcher includes: 1) trigger questions; 2) completeness of assessment components; 3) assessment deadline; and 4) teacher objectivity.

The manifestation of the teacher's needs into the system is also accompanied by other features in the system. For example, referring to the first problem: the unstructured student thesis in the resulting article, related to the completeness of the thesis or argument as the main element in an article. This argument should be structured by a thesis supported by data (Toulmin, 2003). In fact, complete arguments are not only supported by data that are in line with the thesis, but can also be followed by data that invalidate the thesis. For this reason, in addition to realizing the trigger questions, the assessment system includes a thesis statement and data as part of the content assessment carried out by human rater. Then, the second problems: completeness of assessment components, Brown (2003) which mentions written components, such as communication, organization, discourse, syntax, vocabulary, and mechanics, are applied in this system, both for machine rater and human rater.

In the last problem related to teacher objectivity, the system built, as stated in the introduction, is a computer-based test (CBT) with machine rater and human rater (Brown, 2003). The following is an assessment design for writing computer-based opinion articles in Indonesian language according to the needs of Indonesian language teachers, which is in the form of CBT (Figure 2).

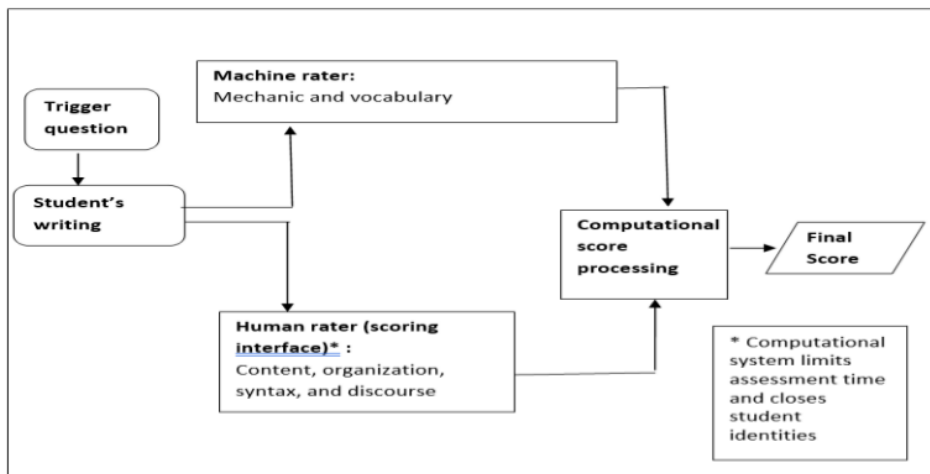


Figure 2. Computer based writing assessment design

In Figure 2, the way the system works is initiated by students who input writing in the form of opinion articles based on trigger questions that have been previously inputted by the teacher. Students write opinion articles in separate columns. Upon completion, the system will conduct an assessment with two raters: machine rater and human rater (Indonesian language teacher). Machine raters apply preprocessing steps in natural language processing (NLP) to assess mechanic and vocabulary (Gelbukh, 2014; Hyland, 2009). These steps are input text editor, sentence segmentation, case folding, tokenizing, stemming, and stopwords removal.

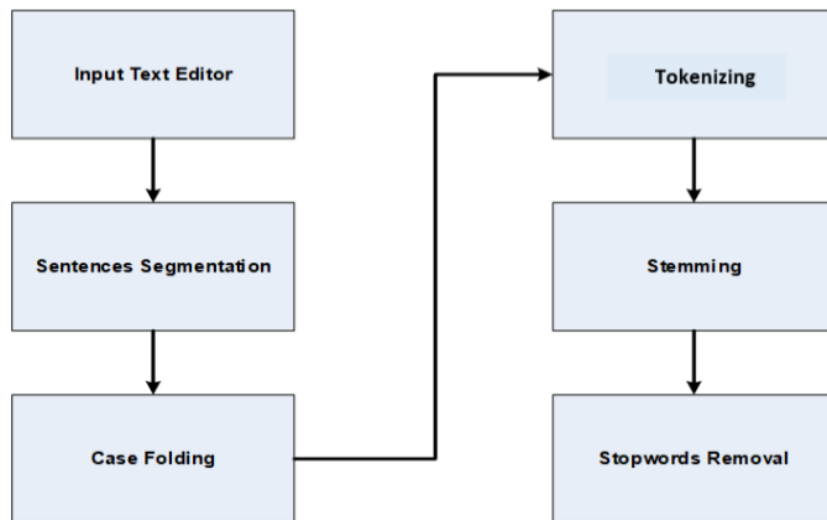


Figure 3. NLP preprocessing stage for the assessment of mechanics and vocabulary

Sentence segmentation is the process of separating sentences and case folding is the process of converting all letters to lowercase, removing irrelevant numbers, punctuation marks, and empty characters. Tokenizing is the process of separating text into chunks known as tokens. Tokens can be words, numbers, symbols, punctuation marks and other important entities. Tokenizing the word means to separate the words in a sentence. Tokenizing is done after casefolding so that the sentence does not contain punctuation, capital letters, empty characters, and

unnecessary numbers. Stemming is the process of changing affixed words into basic words, filtering or stopwords removal is the stage of taking important words from the token results using a stoplist algorithm (removing less important words) or a wordlist (storing important words) (Nugroho, 2019). In this study, the preprocessing steps in NLP (Figure 3) were used as the basis for calculating the number of words produced by students as well as detecting the accuracy of using punctuation marks and word formation.

Meanwhile, human rater assess content, organization, syntax, and discourse using interface assessments. This interface assessment was adapted from a writing test assessment conducted by the ¹² Test of English for International Communication (TOEIC) organized by the Educational Testing Service (ETS) (Everson, 2010). The content consists of the thesis and data, sebagaimana telah dijelaskan di atas bahwa tesis dan data yang mendukungnya merupakan syarat utama isi dari artikel opini (Toulmin, 2003). The organizational component is prepared for teachers to assess the opening paragraphs, content paragraphs, and closing paragraphs (Deane, 2011). The discourse component consists of topic cohesion and continuity (Halliday, 2014), while a syntactic component is available to assess sentence structure (Moeliono, 2017). Each component assessed by human rater is presented in a template that allows the teacher to work fairly and objectively because the assessment components are available in ranges and weights (Brown, 2003) and student identities have been closed.

Assessment by machines will automatically be completed earlier when compared to human raters. However, processing to the final grade is carried out simultaneously after the human rater has finished assessing. Here there is a time limitation established by the system, which limits the human rater to complete the assessment within a certain time. It overcomes the constraints based on needs analysis, where the teacher is unable to complete the manual writing assessment on time due to being interrupted by other activities.

Conclusion

The design of a computer-based writing assessment in Indonesian specifically built to assess opinion articles of grade XII high school students seeks to answer the needs of Indonesian language teachers based on a needs analysis conducted by researchers. The constraints encountered in manual assessment are attempted to be built into this computational system. The inability of students to produce a good thesis if the teacher does not implement trigger questions, is attempted in the system by inputting trigger questions by the teacher before the test begins. The assessment component that is sometimes incomplete in manual assessment is overcome by building a more detailed assessment component so that the teacher can fully assess the students' writing and the student's performance can be seen in its entirety.

The difficulty of the teacher in completing the assessment in a timely manner was overcome by the limitation of the time for the assessment and the machine rater to assess some components: mechanics and vocabulary. The difficulty of teachers in being fair and objective is overcome by the computational system by covering the names of students when the teacher is doing the assessment. The system also provides components and assessment parameters that must be filled in by the teacher as a whole.

This computer-based writing assessment system in Indonesian language can be used for the teaching and learning process, both at school (offline and online) and at home for each student with a network connection so that they can directly connect with the teacher. This system can be adapted for the assessment of other language skills, such as reading, and can also be adapted for the implementation of large-scale tests, such as the Indonesian language proficiency test.

Acknowledgements

I sincerely thank Prof. Dr. Emzir, M.Pd., Dr. Miftahulhairah Anwar, M.Hum. and Dr. Ninuk Lustyantje, M.Pd. at Universitas Negeri Jakarta, Indonesia, who provided advice on this study. My gratitude also goes to the Ministry of Education and Culture of the Republic of Indonesia and the Language Development and Fostering Agency for providing study scholarship support.

References

- Anak Agung Putri Ratna et. al. (2015). *SIMPLE-O, the Essay Grading System for Indonesian Language Using LSA Method with Multi-Level Keywords*. The Asian Conference on Society, Education & Technology 2015. Official Conference Proceedings.
- Borg, Walter L. et.al. (2003). *Educational Research: An Introduction*. USA: Pearson Education, Inc.
- Brown, Douglas H. (2003). *Language Assessment Principles and Classroom Practices*. California: Longman.
- Creswell, John W. & J. David Creswell. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. California: Sage Publications, Inc.
- Deane, Paul. (2011). *Writing Assessment and Cognition*. ETS, Princeton, New Jersey.
<http://www.ets.org/>
- Everson, Philip and Susan Hines. (2010). *Compendium Study: How ETS Scores the TOEIC Speaking and Writing Test Responses*. ETS.
- Gelbukh, Alexander (ed.). (2011). *Computational Linguistics and Intelligent Text Processing*. New York: Springer.
- Halliday, M.A.K. Revised by Christian M.I.M. (2014). *Halliday's Introduction to Functional Grammar*. London and New York: Routledge.

Hyland, K. (2009). *Teaching and researching writing. 2nd edition.*
 10 [https://doi.org/10.1016/S0889-4906\(03\)00034-6](https://doi.org/10.1016/S0889-4906(03)00034-6)

Luh Anik. (2019). "Aplikasi Penyuntingan Ejaan Bahasa Indonesia (Sipebi) 20 Untuk Meningkatkan Mutu Penggunaan Bahasa Indonesia di Lingkungan Badan Pengembangan Bahasa dan Perbukuan". 9 (Laporan Proyek Perubahan Pendidikan Dan Pelatihan Kepemimpinan Tingkat IV).

Lyons, Liz Hamp. 13 (2014). *Writing Assessment in Global Context. Research in The Teaching of English Journal. Volume 48.*

Moeliono, Anton M. 15 (2017). *Tata Bahasa Baku Bahasa Indonesia.* Jakarta: Badan Pengembangan dan Pembinaan Bahasa.

2 Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 24 Tahun 2016 tentang Kompetensi Inti dan Kompetensi Dasar Pelajaran pada Kurikulum 2013 pada Pendidikan Dasar dan Pendidikan Menengah.

Shermis et.al. 5 (2010). *Automated Essay Scoring: Writing Assessment and Instruction. The University of Akron.*

19 Toulmin, S. (2003). *The Uses of Argument. Philosophical Studies.* Cambridge University Press. <https://doi.org/10.5840/philstudies19599042>.

<https://medium.com/@ksnugroho/dasar-text-preprocessing-dengan-python-a4fa52608ffe>.

(accessed: 28 June 2020).

ORIGINALITY REPORT

12%

SIMILARITY INDEX

11%

INTERNET SOURCES

6%

PUBLICATIONS

6%

STUDENT PAPERS

PRIMARY SOURCES

1	semnaskbsp.ums.ac.id Internet Source	1%
2	www.slideshare.net Internet Source	1%
3	ejournal.unsri.ac.id Internet Source	1%
4	Anak Agung Putri Ratna, Budi Selamat Raharjo, Prima Dewi Purnamasari, Randy Sanjaya. "Automatic essay grading system with latent semantic analysis and learning vector quantization", Proceedings of the 3rd International Conference on Communication and Information Processing - ICCIP '17, 2017 Publication	1%
5	mafiadoc.com Internet Source	1%
6	janapflaeging.files.wordpress.com Internet Source	1%
7	"The Handbook of Technology and Second	

Language Teaching and Learning", Wiley, 2017

Publication

1%

8

Submitted to Goldsmiths' College

Student Paper

1%

9

es.scribd.com

Internet Source

1%

10

etheses.bham.ac.uk

Internet Source

1%

11

Submitted to Universitas Negeri Jakarta

Student Paper

1%

12

Submitted to essex

Student Paper

<1%

13

Submitted to Staffordshire University

Student Paper

<1%

14

www.firstpointjapan.com

Internet Source

<1%

15

ojs.badanbahasa.kemdikbud.go.id

Internet Source

<1%

16

en.wikipedia.org

Internet Source

<1%

17

repository.upi.edu

Internet Source

<1%

18

papers.iafor.org

Internet Source

<1%

19 Rama Manor. "The quality of Second-Language Writing (Hebrew) among Arab students in Israel", *Multilingual Education*, 2016 <1%
Publication

20 repositories.kemdikbud.go.id <1%
Internet Source

21 ejournal.unib.ac.id <1%
Internet Source

22 moam.info <1%
Internet Source

23 digilib.stmik-banjarbaru.ac.id <1%
Internet Source

Exclude quotes Off

Exclude matches Off

Exclude bibliography Off