CHAPTER III
RESEARCH METHODOLOGY

3.1 Research Approach and Design

This study employs a descriptive qualitative approach with embedded quantitative lexical analysis. The main objective is to describe and interpret the lexical complexity found in students' descriptive writing, while also exploring factors that may influence their lexical choices. The primary data consists of students' authentic writing samples, which are analyzed using the Lexical Complexity Analyzer (LCA) to obtain quantitative measures such as lexical density, lexical variation, and lexical sophistication. However, no inferential statistical tests are conducted, and the numerical results serve as descriptive support for qualitative interpretation.

In addition to analyzing the students' texts, the study also incorporates data from student questionnaires. These questionnaires are designed to examine students' reading habits, perceptions, and writing strategies. This supporting data is analyzed qualitatively to help explain possible factors contributing to the lexical patterns observed in their writing.

3.2 Research Subject and Setting

The study was conducted at SMA Muhammadiyah X Surabaya, a senior high school in Indonesia where English is taught as a foreign language. Eleventh-grade students were selected as participants because they are in a critical stage of cognitive and linguistic development, as described by Vygotsky's Zone of Proximal Development (ZPD). At this level, students have acquired sufficient

foundational writing skills and are cognitively ready to develop more advanced academic literacy.

In this study, one class of female students was selected using purposive sampling. This non-probability sampling technique allows the researcher to select participants based on specific considerations relevant to the research objectives. The selected class was deliberately chosen because it includes students with special educational needs (inclusive students), whose presence provides a more diverse linguistic profile for the study. The inclusion of these students is considered important, as it may influence the lexical complexity of classroom writing tasks, providing richer data for analysis.

This purposive approach also considered practical factors such as the class's availability, accessibility, and alignment with the research focus. Using purposive sampling ensures that the sample represents the specific characteristics needed for the study, particularly regarding inclusive education and its potential impact on students' language use.

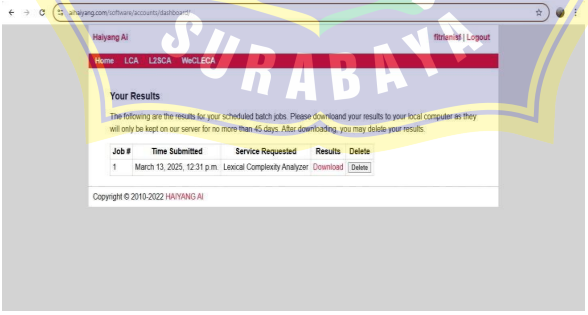

Figure 1 Lexical Complexity Analyzer WEB

Furthermore, based on the Indonesian senior high school English curriculum, descriptive texts are explicitly taught and emphasized in eleventh grade, making this level appropriate for examining students' lexical complexity in descriptive writing. The students previously completed descriptive writing assignments, which were available in digital format, were repurposed as research data to facilitate analysis using the Lexical Complexity Analyzer (LCA).

3.3 Research Instrument

In this study, the writers employed three key research tools to gather and analyze data, each contributing to a deeper understanding of lexical complexity in students' writing.

The main instrument in this research a writing task based on one same topic to keep the data consistent. All students were asked to write a descriptive text using the same topic. It is only used to make sure that the writing task is done fairly and consistently by all students. The main focus of this study to examine the lexical patterns in students' writing.

To support the evaluation of students' descriptive texts, a writing assessment rubric was utilized in this study. The rubric assesses general aspects of writing, including content,

organization, grammar, vocabulary, and mechanics, providing a structured framework to evaluate the overall quality of students' written work. According to Brown (2004), rubrics play a crucial role in standardizing the assessment of writing, as they offer clear and consistent criteria to measure students' performance

objectively. This rubric was adapted and validated through expert judgment prior to its implementation, ensuring its appropriateness and reliability for evaluating student writing, as recommended by O'Malley and Pierce (1996). The use of the rubric complements the lexical analysis by offering a qualitative perspective on students' writing performance beyond quantitative lexical measures

The second is used the Lexical Complexity Analyzer (LCA), a web tool created by *Haiyang Ai* which can accessed in https://aihaiyang.com/software/lca/batch/.

Last instrument, the researchers employed a structured questionnaire designed to investigate students' attitudes, beliefs, and behaviors toward reading and writing in English. about their exposure to English input, their feelings toward writing tasks, and the strategies they used for vocabulary development.

By combining these three tools, the researchers were able to obtain a comprehensive view of both the linguistic output (through the analysis of writing samples) and the underlying behavioral factors (through the questionnaire) that may influence lexical complexity in students' writing.

3.4 Data Collection Technique

This study employed two main techniques for data collection: text analysis using the Lexical Complexity Analyzer (LCA) and the distribution of questionnaires. The written data were collected from expository texts produced by eleventh-grade students, which were then analyzed using the LCA tool accessed through its

official website. This tool is use to measure selected indicators of lexical complexity. In addition to text analysis, questionnaires were administered to gather supporting information about students' writing habits and the factors that may influence their lexical choices.

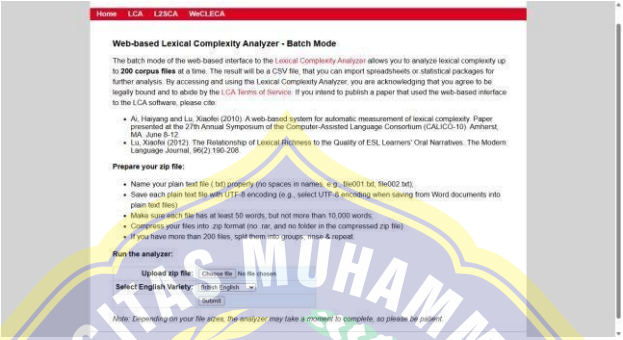

Figure 2 Tutorial LCA

3.4.1 Lexical Analyzer (LCA)

This tool, LCA works with *.txt* files that contain only plain paragraphs without images, graphs, tables, figures, references, titles, or subtitles. The text must also follow American spelling, which is checked through a scanning process. The Web-based LCA provides numerical results that represent different aspects of lexical complexity, including lexical density, lexical sophistication, and lexical variation.

If you follow the tutorial correctly, you can upload the document into the LCA tool in the proper format. Once processed, the results was appeared on the website and can be downloaded as an *.xlsx* file.

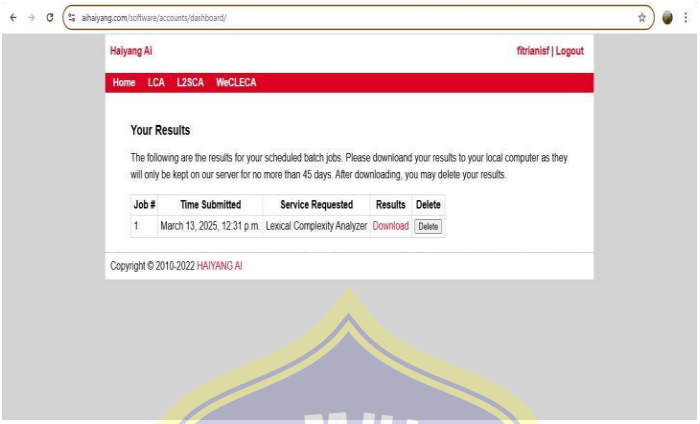

Figure 3 Tutorial LCA

In this study, lexical complexity analysis conduct using the Lexical Complexity Analyzer (LCA), but not all indicators available in the tool were utilized. The selection of indicators made based on their relevance and suitability for the research objectives, which focus on measuring lexical complexity in expository texts written by eleventh-grade students. The indicators used include *Lexical Density (LD),* Type-Token Ratio (TTR), Number of Different Words (NDW), and Lexical Sophistication (LS1), each of which provides insight into the depth and variation of vocabulary in the texts. Additionally, Vocabulary Sophistication (VS1) used as an optional indicator, selected based on further considerations to examine the students' vocabulary proficiency in the context of their writing. The choice of these specific indicators aims to focus on aspects of lexical complexity that can provide a deeper understanding of the quality and diversity of vocabulary used by the students, without the need to analyze all the indicators available in the LCA. This approach allows for a more

focused and in-depth analysis of the factors influencing lexical complexity in student writing.

3.4.2 Questionnaire

To complement the lexical analysis using the Lexical Complexity Analyzer (LCA), this study employed a questionnaire to explore factors that might influence students' lexical complexity in writing. aiming to examine students' beliefs, feelings, and behaviors related to reading and writing in English. Additional items were also constructed to assess students' lexical awareness, exposure to descriptive texts, and writing strategies. As noted by Hesmatantya (2023), such differentiated assessments can help address the diverse learning needs within inclusive classrooms by identifying both cognitive and affective barriers to vocabulary acquisition.

Before distribution, the questionnaire was validated through expert judgment to ensure its content validity, clarity, and appropriateness for the target participants. An expert validator reviewed the questionnaire items, focusing on the relevance of each statement to the research objectives and the linguistic suitability for high school students. Based on their feedback, minor revisions were made to improve item wording and coherence.

The questionnaire consisted of 21 statements using a 5-point Likert scale, ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). These items were grouped into several categories: cognitive, affective, and behavioral components of attitude; exposure to English descriptive texts; vocabulary awareness; perceptions of writing tasks; and specific lexical strategies. The responses were analyzed to provide contextual insight into students' lexical development with

interpretation ranges commonly adopted in attitude studies (Boone & Boone, 2012), where scores from 3.50 to 5.00 are interpreted as indicating a positive attitude. Therefore, a mean score of 4.13 in the cognitive component reflects a strong belief among students about the value of reading for writing improvement.

3.5 Data Analysis

The data in this study were analyzed using both quantitative and qualitative approaches. The primary data, consisting of descriptive texts written by eleventh-grade students, were analyzed using the Lexical Complexity Analyzer (LCA) an online tool designed to measure various dimensions of lexical complexity. The aspects analyzed include lexical diversity, lexical sophistication, and lexical density. The researcher accessed the LCA tool through its official website and uploaded each student's text to obtain the lexical complexity scores. The results were then interpreted to determine the overall level of lexical complexity in students' writings.

In addition to text analysis, supporting data were collected through a questionnaire consisting of 21 Likert-scale items. The questionnaire responses were tabulated and categorized based on their corresponding theoretical components: the Input Hypothesis and the Tripartite Model of Attitude. Descriptive statistics were used to identify trends and patterns in students' attitudes, exposure to reading materials, lexical awareness, and writing strategies. These findings were used to support and explain the lexical patterns observed in the written texts, providing a more comprehensive understanding of the factors that may contribute to lexical complexity in students' writing.