

# **Favorite Vocabulary in Students' Thesis and Dissertation of S-2 and S-3 Study Programs Language and Literature Education**

**Maria Mintowati, Kisyani-Laksono, Ahmad Bayu Prastyo**

Pascasarjana Universitas Negeri Surabaya

mintowati@unesa.ac.id, kisyani@unesa.ac.id, ahmad.19049@mhs.unesa.ac.id

**Abstract:** One of the requirements for graduation as a master or doctor is a graduate student writing a final paper in the form of a thesis or dissertation. Through a relatively long writing process to realize the written work, ideas are expressed through the choice of words and their arrangement in sentences to paragraphs and discourse. The wealth of written vocabulary is one indicator of mastering one's writing skills. In this paper, favorite vocabulary is discussed (high frequency of occurrence) in the thesis (T-1, T-2, T-3, T-4) and dissertation (D-1, D-2, D-3, D-4). Data is collected by documentation techniques, then processed with "Aplikasi Lema." The results of this study are found three favorite words in T-1: "Banyuwangi" (176); "Iqbal" (153); "Singodimay" (109); T-2: "supplement" (282); "Trefinger" (150); T-3: "kapata" (322); "Lounusa" (228); "Latu" (201); T-4: "Tejo" (137); "Sujiwo" (112); "Indonesia" (49). Next, three favorite words in D-1: "crite" (293); "Production" (263); "RPS" (257); D-2: "Germany" (1072); "neurodidactic" (451); "Deutsch" (218); D-3: "dp" (945); "berian" (531); "Gorom" (289); D-4: "differentials" (847); "evidence" (610); "entry" (595). The use of written vocabulary in the thesis and dissertation is dominated by words in the field of science. This finding is very much different from research on Indonesian students' junior high school textbooks (Kisyani, et al., 2017) and Indonesian high school (Kisyani, et al., 2018). In these two studies, the favorite word is dominated by the conjunction "yang", "dan", "serta". This is due to the data sources used in this study, namely scientific work in the form of theses and dissertations, while previous studies used the books of middle and high school students, the majority of which were discourse.

**Keywords:** *favorite vocabularies, thesis, dissertation*

## **Introduction**

One of the language elements that have an essential role for language learners is vocabulary. It is said so because through mastering the vocabulary of a language, they have a vocabulary (Blachowicz and Fisher, 2004). Through this, they can understand the meaning and increase knowledge about the language learned and use it in a conversation or writing. From that explanation, it is natural that vocabulary ownership, learners expressed rich vocabulary (cf. **Laufer, 2010; Lupescu, 1993**).

In completing an education level, besides taking several courses, language learners should write scientific papers. For each level in tertiary education (Strata 1, Strata 2, and Strata 3) are named as thesis undergraduate, thesis, and dissertation. To realize this, they are required to master academic material and the ability to express ideas until scientific work is born. In addition to the depth of the material, the actuality of the research topic, the accuracy of the research method, the use of language is one of the evaluation criteria of scientific work (Pedoman Penulisan Tesis dan Disertasi Pascasarjana Unesa, 2016).

Master and doctoral students of the Postgraduate Language and Literature Education Study Program at Surabaya State University are also required to write scientific papers in the form of a thesis or a dissertation. In pouring their ideas, they use varied written vocabulary, both in the number of words, entries, and favorite words (words with high frequency). The written

vocabulary they use in these theses and dissertations is interesting to note, especially words with the high frequency of use.

In Kisyani, et al. (2017), the favorite vocabulary in student books for Indonesian junior high school subjects is task words, namely "yang", "dan", and "serta." This is also evident in the research of Kisyani, et al. (2018) with a data source in the form of a book for Indonesian high school students. What is interesting is that the findings are not in the form of charged words, for example, nouns, verbs, and/or adjectives, but task words in the form of conjunctions. Starting from this, this paper will discuss any favorite written vocabulary and how it appears in theses and dissertations, especially the work of S-2 and S-3 students of Language and Literature Education Study Program, Postgraduate, Surabaya State University. The determination of this data source is based on the assumption that these students have more linguistic knowledge and skills when compared to students from other study programs available at Postgraduate Unesa. This is considered logical considering that students from both study programs have and are studying language with all aspects of it.

## **Method**

This research is a descriptive study because it aims to describe the favorite vocabulary in the thesis and dissertation of students of the Language and Literature Education Study Program at the State University of Surabaya. The data source consists of five theses and five dissertations, each coded T-1, T-2, T-3, T-4, T-5, and D-1, D-2, D-3, D-4, D-5. Data in the form of written vocabulary which frequency of occurrence most often (favorite vocabulary in the data source. Data is collected using the application "lema" (Kisyani, et al. 2018). The application is developed with the help of PT Andromedia. Form of application in the first year (2017) in the form of tools This application can calculate entries in less than one minute The results of the calculation consist of three sheets in the Excel program, namely: summary, recapitulation, report, the following are the rules and procedures in the application.

- 1) The maximum amount of data that can be calculated is 10 MB.
- 2) The amount of data is then uploaded in the application with the following steps:
  - a. The application is opened in the main file
  - b. The initial display shows pdf and database scans.
  - c. Data in pdf format is uploaded with the data upload menu.
  - d. In a matter of less than one minute, a "pdf rendering was successful."
  - e. Short results immediately posted
  - f. Click the download menu that appears.
  - g. The downloaded data is presented in the form of an excel file.
  - h. Database entries can be added or subtracted, especially if there are starred results..

The raw material used is the list of entries, re-words, and affixes as a database (cf. Inayatillah, 2019). Before becoming a database, the list is checked first with the help of the excel, sort, pivot table program. Then affix analysis is done so that it can become a database by giving examples of initial letters for all words that get affixes. On the other hand, affixes are included in the list of exceptions, so they are not recognized as entries.

The resulting database is then entered into the application. The entry application consists of: backend, bin, data folder, UI. Besides, there are also install.txt, main.exe.main.go, and README.md. The name "play" is what is clicked to open the application. When the application

is open, there will be a display at the top left of the PDF and Database Scan. Underneath, there is a Scan PDF command and a box containing the click browse button to scan the document. At the bottom, there is a "File:" button and a "Browse PDF" button. In the middle, there are the most words, and the correct ten entries are listed.

If the "browse PDF" button is clicked, it will point to the file to be selected. The file is available in PDF format. If the file has been found, then click "open," and the display will return to the original menu with the part of the file that is already filled. On the right side of the PDF browser, a blue "PDF process" button will appear. If the blue button is clicked, less than one minute, a "Render PDF was successful" notification appears and there is a green "OK" button on it. If the "OK" button is clicked, the number of words will be filled, and the ten most entries will be filled.

At the bottom left, there is an article to download the export results. Click the green "download" button. If the green button is clicked, the save as display will appear, what file will be named (please write), and click the "save" button. Then the screen will return to the initial menu and there is a successful download, and there is an "ok" button to click. The results report is in the excel file that was downloaded earlier. With this embodiment process, the resulting application is able to calculate entries in less than one minute. The calculation results consist of three sheets in the Excel program, namely: summary, summary, report.

**Results and Discussion**

By using the application "Lema" (Kisyani, et al., 2018), the files of five theses and five dissertations that have been converted to pdf are calculated by the number of pages, the number of words, and the number of entries. The calculation results are as follows.

**Table 1.** Thesis Identification Result Result Thesis Postgraduate Study Program of Language and Literature at Unesa Postgraduate Program

No	Aspects	T-1	T-2	T-3	T-4	T-5
1	Number of pages	122	136	216	88	118
2	Number of words	22976	20760	41590	15204	19561
3	Number of entries	1377	743	1727	1063	622

From the table, it appears that the thesis with the thickest number of pages is T-3, then followed by T-2, T-1, T-5, and T-4. Then, the most number of words is in T-3, then followed by T-1, T-2, T-5, and T-4. Based on the two calculation results, it is found that between the number of pages and the number of words is directly proportional. This is different from the number of entries in the five theses. Based on the number of entries, the most entries are in T-3, then followed by T-1, T-4, T-2, and T-5. What remains consistent is between the number of pages: the number of words: the number of entries in T-3. That is, the number of pages, words, and entries are equally numerous. This was also found in T-5 and T-4, whereas in T-2 and T-1, this did not happen.

The following are the results of counting the number of pages, the number of words, and the number of entries in the five dissertations of Unesa Language and Literature Education Study Program students.

**Table 2.** Dissertation Vocabulary Identification Results Postgraduate Doctoral Program Student of Language and Literature at Unesa Postgraduate

No	Aspects	D-1	D-2	D-3	D-4	D-5
1	Number of pages	294	284	264	338	426
2	Number of words	57193	48196	41968	68509	88848
3	Number of entries	1797	1867	2272	2003	2477

In table 2, the highest number of pages is on D-5, followed by D-4, D-1, D-2, and D-3, respectively. The highest number of words is in D-5, then followed by D-4, D-1, D-2, and D-3. Furthermore, the highest number of entries is in D-5, then followed by D-3, D-4, D-2, and D-1. Based on the results of data calculation with the application "LEMA" it can be stated that there are results directly proportional between the number of pages and the number of words. This is proven. However, finding the number of entries when compared to the number of pages as well as the number of words, this is not directly proportional. In table 2, the results show a direct proportion between the number of pages, the number of words, and the number of entries found in D-5 and D-2. The reverse findings are proven by the number of pages, the number of words, and the number of entries that are not directly proportional to D-4, D-3, and D-4.

Returning to the main problem in this paper is what favorite word appears in five theses and five dissertations as sources of research data. Based on the work of the application, "Lema" on five theses and five dissertations found ten favorite written vocabulary as presented in the table as follows.

**Table 3.** Favorite Vocabulary in Thesis Unesa Language and Literature Study Student

No	T-1	Number of entries	T-2	Number of entries	T-3	Number of entries	T-4	Number of entries	T-5	Number of entries
1	Banyuwangi	176	Suplemen	282	Kapata	322	tejo	137	Discovery	250
2	Iqbal	153	Treffing-er	150	lounusa	228	sujiwo	112	Billfath	69
3	Singodimayan	109	valid	146	Latu	201	Indonesia	49	Validator	60
4	Pki	72	Validator	123	Sopo	200	Rahwana	40	Rps	57
5	Rlin	70	(PE)rsentase	38	Tamilouw	188	Rahva-yana	39	Kumpul	50
6	Hasn	58	posttest	32	amalatu	177	lala	38	Praktis	42
7	Santet	58	(PE)roleh	29	Trum	167	Tatan	24	Terlak-sana	41

No	T-1	Number of entries	T-2	Number of entries	T-3	Number of entries	T-4	Number of entries	T-5	Number of entries
8	Negari	49	rancang	25	Hutumuri	165	Sinta	19	Rancang	39
9	Budoyo	49	pretest	24	Bongso	135	Mua	18	Peroleh	34
10	Jurji	49	respon	24	Sori	125	Kuasa	18	Tuntas	34

T-1 Les; T-2 Sin; T-3 Kes; T-4 Nur; T-5 Roh,

On T-1, three favorite writing vocabularies are (1) "Banyuwangi", (2) "Iqbal", and (3) "Singodimayan". T-1 titled "Expression of Cultural Adaptation in the Novel Santet Gandrung Veil and Niti Negari Bala Abangan by Hasnan Singodimayan". Favorite words "Banyuwangi" is a place setting contained in literary works. The "Iqbal" vocabulary is the main character, while the "Singodimayan" vocabulary is the name of the novel's author. Based on that, favorite vocabulary arises based on the tendency of research variables.

On T-2, three favorite vocabulary words are (1) "Supplements", (2) "Treffinger", and (3) "Valid". It has the same tendency on T-1, namely favorite vocabulary appears on the research variable. Likewise in T-3, T-4, and T-5, all three have the same level of emergence of favorite vocabulary. This shows that Unesa's Indonesian language and literature education student thesis has a good vocabulary.

**Table 4.** Favorite Vocabulary in the Dissertation of Unesa Language and Literature Study Program Students

No	D-1	Number of entries	D2	Number of entries	D3	Number of entries	D4	Number of entries	D-5	Number of entries
1	Crite	293	Jerman	1072	Dp	945	Diferensia	847	item	68
2	Producti-n	263	Neurodidaktik	451	Berian	531	Evidensi	610	kukur	62
3	Rps	257	Deutsch	218	Groom	289	Lema	595	rtl	57
4	Prancis	210	Cefr	96	Glos	184	Konsep	557	test	42
5	Integrasi	170	Zertifikat	93	Pns	174	End	168	signifi	32
6	Niveau	170	Lulus	58	Pan	163	Pragmatik	159	sni	26
7	Validator	120	Capai	54	Dialektologi	161	Circulus	152	pre	25
8	Rancang	109	Praktis	48	Leksikal	131	Define	128	Product	25
9	Capai	104	Kampung	43	Konservatif	131	Padan	96	Pustaka	24
10	Kkni	86	Rps	39	Hitung	93	Pragmasemantik	88	bubut	22

D-1 Gh; D-2 Krt; D-3 Rml; D-4 Aml; D-5 Arn.

On D-1, three favorite writing vocabularies are (1) "Crite", (2) "Production", and (3) "Rps". Favorite vocabulary "Crite" is part of the Production course. The "Production" vocabulary is the subject being taught, while the "Rps" vocabulary is the instrument used by the writer. Based on that, favorite vocabulary arises based on research variables.

On D-2, three favorite vocabulary words are (1) "Jerman", (2) "Neurodidaktik", and (3) "Deutch". It has the same tendency in D-1, namely, favorite vocabulary appears on the research variable. Likewise, in D-3, D-4, and D-5, all three have the same level of emergence of favorite vocabulary. This shows that Unesa's Indonesian language and literature education student thesis has a useful vocabulary.

## **Conclusions**

Favorite writing vocabulary in the thesis and dissertation is dominated by words in the field of science. Favorite KT appears based on the research title variable. This is different from the reading vocabulary in student books for Indonesian subjects, both junior high and high school level. Favorite vocabulary in the two different level student books is dominated by prepositions and particles. Thus, the favorite KT in the Unesa Indonesian Language and Literature Education students' theses and dissertations demonstrates mastery of core vocabulary, according to the topics in the thesis and dissertation they examined.

## **References**

- Blachowicz, Camile L.Z. and Peter Fisher. (2004). "Vocabulary Lessons". *Journal of Educational Leadership*, March 2004, Vol. 6 (6).
- Carrol, Devies, dan Richman. (1971). "The American Herritage Word Frequency Book".  
in [http://en.wikipedia.org/wiki/Word\\_lists\\_by\\_frequency](http://en.wikipedia.org/wiki/Word_lists_by_frequency). Diunduh 12 Februari 2016.
- Inayatillah, F. et all. (2019, December). Entry application software to identify the development of reading and writing vocabulary. In *Journal of Physics: Conference Series* (Vol. 1339, No. 1, p. 012117). IOP Publishing.
- Kisyani, et all. (2017). "Kosakata Baca dan Kosakata Tulis Siswa SMP". Laporan Penelitian. Surabaya: LPPM Universitas Negeri Surabaya.
- Kisyani. (2018). "Pembelajaran Bahasa Dan Sastra Indonesia Berbasis Kehidupan". Malang: UM.
- Kisyani, et all. (2019, December). Vocabulary Learning on Junior High School Students' Textbook. In *Social Sciences, Humanities and Education Conference (SoSHEC 2019)*. Atlantis Press.
- Kisyani, et all. (2018). ). "Kosakata Baca dan Kosakata Tulis Siswa SMA". Laporan Penelitian. Surabaya: LPPM Universitas Negeri Surabaya.
- Kisyani, et all. (2018). "The Development of Vocabulary in Bahasa Indonesia Textbook for Senior High School" dalam *Advances in Social Science, Education and Humanities Research, volume 222 2nd Social Sciences, Humanities, and Education Conference (SoSHEC 2018)* (Atlantis Press).
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). "Lexical Threshold Revisited: Lexical Text Coverage, Learners' Vocabulary Size and Reading Comprehension". *Reading in a foreign language*, 22(1), 15-30.
- Lupescu, S., & Day, R. R. (1993). *Reading, dictionaries, and vocabulary learning*. Language

learning, 43(2), 263-279.

Mintowati, M. et all. (2018, December). Regional and Foreign Vocabulary in the Junior High School Bahasa Indonesia Student Books. In 2nd International Conference on Education Innovation (ICEI 2018). Atlantis Press.

Zuchdi, Darmiyati. 1997. "KB Bahasa Indonesia untuk Sekolah Dasar di Indonesia."

Yogyakarta: FPBS IKIP Yogyakarta.